

| | | | | |
|-----------------------|---|---|-----------------|---------------------------------------|
| FT | MOD_RES | 375 | 375 | Phosphothreonine (by CK2) (in vitro). |
| FT | MOD_RES | 379 | 379 | Phosphoserine (by CK2) (in vitro). |
| SQ | SEQUENCE | 493 AA; | 55111 MW; | 9694CD302BEBBEDE CRC64; |
| Query Match | | 60.3%; | Score 79; | DB 1; Length 493; |
| Best Local Similarity | | 58.3%; | Pred. No. 0.02; | 5; Indels 0; Gaps 0; |
| Matches | 14; | Conservative | 5; Mismatches | |
| Qy | 3 | REESSEYMAADEYNRLKQVKQPA 26 | | |
| | | | | |
| Db | 442 | RKDSSEYRTVADKYNRLKEIKSSA 465 | | |
| | | | | |
| RESULT 13 | | | | |
| Q6NRR0 | | | | |
| ID | Q6NRR0 | PRELIMINARY; | PRT; | 497 AA. |
| AC | Q6NRR0; | | | |
| DT | 05-JUL-2004 | (TrEMBLrel. 27, Created) | | |
| DT | 05-JUL-2004 | (TrEMBLrel. 27, Last sequence update) | | |
| DE | 05-JUL-2004 | (TrEMBLrel. 27, Last annotation update) | | |
| DE | MG82312 | protein. | | |
| GN | Name=MG82312; | | | |
| OC | Xenopus laevis (African clawed frog). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; | | | |
| OC | Xenopodinae; Xenopus. | | | |
| NCBI_TaxID=8355; | | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | TISSUE=Kidney; | | | |
| RX | MEDLINE=22398257; PubMed=12477932; | | | |
| RA | Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G., | | | |
| RA | Rahs S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., | | | |
| RA | Richardson D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., | | | |
| RA | Whiting J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A., | | | |
| RA | Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., | | | |
| RA | Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., | | | |
| RA | Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E., | | | |
| RA | Jones S.J., Marra M.A. | | | |
| RT | "Generation and initial analysis of more than 15,000 full-length human | | | |
| RT | and mouse cDNA sequences." | | | |
| RL | Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002). | | | |
| RL | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | TISSUE=Kidney; | | | |
| RX | MEDLINE=22341132; PubMed=12454917; | | | |
| RA | Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W., | | | |
| RA | Richardson P. | | | |
| RT | "Genetic and genomic tools for Xenopus research: The NIH Xenopus | | | |
| RT | initiative." | | | |
| RL | Dev. Dyn. 225:384-391(2002). | | | |
| RL | [3] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | TISSUE=Kidney; | | | |
| RA | Klein S., Strausberg R.; | | | |
| RL | Submitted (May-2004) to the EMBL/GenBank/DBJ databases. | | | |
| DR | EMBL; BC070570; AAH70670.1; | | | |
| DR | InterPro; IPR008253; Marvel. | | | |
| DR | InterPro; IPR002958; Occludin. | | | |
| DR | InterPro; IPR010844; Occludin_ELL. | | | |
| DR | Pfam; PF01284; MARVEL; 1. | | | |
| DR | Pfam; PF07303; Occludin_ELL; 1. | | | |
| DR | PRINTS; PR01258; OCCLUDIN. | | | |
| SQ | SEQUENCE 497 AA; 56202 MW; 269431E14ECD0EAL CRC64; | | | |

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:41:17 ; Search time 19.9552 Seconds
(without alignments)
222.664 Million cell updates/sec

Title: US-10-092-750-20

Perfect score: 330

Sequence: 1 KGIISRLMSVEELKRDHAE.....ARLMSALTAPXRVVXHCXPL 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCUTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|----------------------|
| 1 | 72 | 21.8 | 1135 | 4 | US-09-294-298A-21 |
| 2 | 72 | 21.8 | 1325 | 4 | US-09-294-298A-6 |
| 3 | 72 | 21.8 | 1376 | 4 | US-09-294-298A-4 |
| 4 | 72 | 21.8 | 1423 | 4 | US-09-294-298A-2 |
| 5 | 62.5 | 18.9 | 3562 | 4 | US-09-679-279-14 |
| 6 | 62 | 18.8 | 124 | 4 | US-09-621-976-4246 |
| 7 | 60 | 18.2 | 596 | 4 | US-09-252-991A-32193 |
| 8 | 59.5 | 18.0 | 688 | 4 | US-09-252-991A-25685 |
| 9 | 58 | 17.6 | 494 | 3 | US-09-019-095A-24 |
| 10 | 58 | 17.6 | 519 | 4 | US-09-328-352-7681 |
| 11 | 57.5 | 17.4 | 260 | 4 | US-09-540-236-2236 |
| 12 | 57 | 17.3 | 471 | 3 | US-09-126-420A-20 |
| 13 | 57 | 17.3 | 620 | 4 | US-09-538-092-1285 |
| 14 | 56.5 | 17.1 | 526 | 4 | US-09-198-452A-67 |
| 15 | 56 | 17.0 | 502 | 3 | US-09-499-302A-7 |
| 16 | 56 | 17.0 | 552 | 4 | US-09-252-991A-32288 |
| 17 | 56 | 17.0 | 1233 | 4 | US-09-252-991A-23237 |
| 18 | 55.5 | 16.8 | 612 | 4 | US-09-252-991A-17844 |
| 19 | 55 | 16.7 | 214 | 4 | US-09-252-991A-30067 |
| 20 | 55 | 16.7 | 288 | 4 | US-09-305-856B-2 |
| 21 | 55 | 16.7 | 288 | 5 | PCT-US92-00282-19 |
| 22 | 55 | 16.7 | 533 | 5 | PCT-US92-00282-3 |
| 23 | 54.5 | 16.5 | 323 | 4 | US-09-252-991A-25672 |
| 24 | 54.5 | 16.5 | 333 | 4 | US-09-252-991A-17670 |
| 25 | 54.5 | 16.5 | 810 | 4 | US-09-198-452A-20 |
| 26 | 54.5 | 16.5 | 810 | 4 | US-09-198-452A-1128 |
| 27 | 54 | 16.4 | 180 | 4 | US-09-252-991A-31563 |

ALIGNMENTS

RESULT 1

US-09-294-298A-21

; Sequence 21, Application US/09294298A

; Patent No. 6723838

; GENERAL INFORMATION:

; APPLICANT: HUGANIR, RICHARD L.

; APPLICANT: KIM, GUEHAE

; TITLE OF INVENTION: SIGNAL TRANSDUCING SYNAPTIC MOLECULES AND USES THEREOF

; FILE REFERENCE: 48235/1699

; CURRENT APPLICATION NUMBER: US/09/294,298A

; CURRENT FILING DATE: 1999-04-19

; PRIOR APPLICATION NUMBER: 60/082,690

; PRIOR FILING DATE: 1998-04-22

; PRIOR APPLICATION NUMBER: 60/082,717

; PRIOR FILING DATE: 1998-04-23

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 21

; LENGTH: 1135

; TYPE: PRT

; ORGANISM: Unknown Organism

; FEATURE:

; OTHER INFORMATION: Description of Unknown Organism: mammalian

; OTHER INFORMATION: SYNGAP-C

US-09-294-298A-21

Query Match

Best Local Similarity 21.8%; Score 72; DB 4; Length 1135;

Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY

1 KGIISRLMSVEELKRDHAE 21

1087 KSIIGRLMVEELRDRHPAM 1107

RESULT 2

US-09-294-298A-6

; Sequence 6, Application US/09294298A

; Patent No. 6723838

; GENERAL INFORMATION:

; APPLICANT: HUGANIR, RICHARD L.

; APPLICANT: KIM, GUEHAE

; TITLE OF INVENTION: SIGNAL TRANSDUCING SYNAPTIC MOLECULES AND USES THEREOF

; FILE REFERENCE: 48235/1699

; CURRENT APPLICATION NUMBER: US/09/294,298A

; CURRENT FILING DATE: 1999-04-19

; PRIOR APPLICATION NUMBER: 60/082,690

; PRIOR FILING DATE: 1998-04-22

; PRIOR APPLICATION NUMBER: 60/082,717

; PRIOR FILING DATE: 1998-04-23

Sequence 21202, A
Sequence 186, App
Sequence 20018, A
Sequence 32257, A
Sequence 45785, A
Sequence 28383, A
Sequence 14855, A
Sequence 4, Appli
Sequence 4, Appli
Sequence 7668, Ap
Sequence 31488, A
Sequence 19101, A
Sequence 18060, A
Sequence 28988, A
Sequence 27, Appl
Sequence 17007, A
Sequence 21847, A
Sequence 22465, A

28 54 16.4 389 4 US-09-252-991A-21202
29 53.5 16.2 111 4 US-09-198-452A-186
30 53.5 16.2 154 4 US-09-252-991A-20018
31 53.5 16.2 848 4 US-09-252-991A-32257
32 53 16.1 124 4 US-09-270-767-45785
33 53 16.1 311 4 US-09-252-991A-28383
34 53 16.1 626 4 US-09-248-796A-14855
35 52.5 15.9 292 2 US-08-973-461A-4
36 52.5 15.9 292 3 US-08-648-010-4
37 52.5 15.9 527 4 US-09-489-039A-7668
38 52.5 15.9 669 4 US-09-252-991A-31488
39 52 15.8 347 4 US-09-248-796A-19101
40 52 15.8 461 4 US-09-252-991A-18060
41 52 15.8 1225 4 US-09-252-991A-28988
42 51.5 15.6 107 3 US-09-299-486F-27
43 51.5 15.6 172 4 US-09-252-991A-17007
44 51.5 15.6 351 4 US-09-252-991A-21647
45 51.5 15.6 557 4 US-09-252-991A-22465

| | | |
|---|---|--|
| Query Match | 18.8%; Score 62; DB 4; Length 124; | |
| Best Local Similarity | 48.5%; Pred. No. 0.38; | |
| Matches | 16; Conservative 3; Mismatches 12; Indels 2; Gaps 1; | |
| Qy | 6 RLSVSELKRDHAEMQAGCGLOQTEDHLMRRS 38 | |
| Db | 32 RLVLVLEKLRDHLMSRQCHALQED--MQTRS 62 | |
| | | |
| RESULT 7 | | |
| US-09-252-991A-32193 | | |
| ; Sequence 32193, Application US/09252991A | | |
| ; Patent No. 6551795 | | |
| ; GENERAL INFORMATION: | | |
| ; APPLICANT: Marc J. Rubenfield et al. | | |
| ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS | | |
| ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS | | |
| ; FILE REFERENCE: 107196.136 | | |
| ; CURRENT APPLICATION NUMBER: US/09/252,991A | | |
| ; CURRENT FILING DATE: 1999-02-18 | | |
| ; PRIOR APPLICATION NUMBER: US 60/074,788 | | |
| ; PRIOR FILING DATE: 1998-02-18 | | |
| ; PRIOR APPLICATION NUMBER: US 60/094,190 | | |
| ; PRIOR FILING DATE: 1998-07-27 | | |
| ; NUMBER OF SEQ ID NOS: 33142 | | |
| ; SEQ ID NO 32193 | | |
| ; LENGTH: 596 | | |
| ; TYPE: PRT | | |
| ; ORGANISM: Pseudomonas aeruginosa | | |
| US-09-252-991A-32193 | | |
| Query Match | 18.2%; Score 60; DB 4; Length 596; | |
| Best Local Similarity | 35.1%; Pred. No. 5.9; | |
| Matches | 13; Conservative 3; Mismatches 21; Indels 0; Gaps 0; | |
| Qy | 1 KGIISRLMSVEELKRDHAEMQAGCGLOQTEDHLMRR 37 | |
| Db | 541 KGLAHLALDPHRRHQDHAENHAGAGENPSGHFRGR 577 | |
| | | |
| RESULT 8 | | |
| US-09-252-991A-25685 | | |
| ; Sequence 25685, Application US/09252991A | | |
| ; Patent No. 6551795 | | |
| ; GENERAL INFORMATION: | | |
| ; APPLICANT: Marc J. Rubenfield et al. | | |
| ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS | | |
| ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS | | |
| ; FILE REFERENCE: 107196.136 | | |
| ; CURRENT APPLICATION NUMBER: US/09/252,991A | | |
| ; CURRENT FILING DATE: 1999-02-18 | | |
| ; PRIOR APPLICATION NUMBER: US 60/074,788 | | |
| ; PRIOR FILING DATE: 1998-02-18 | | |
| ; PRIOR APPLICATION NUMBER: US 60/094,190 | | |
| ; PRIOR FILING DATE: 1998-07-27 | | |
| ; NUMBER OF SEQ ID NOS: 33142 | | |
| ; SEQ ID NO 25685 | | |
| ; LENGTH: 688 | | |
| ; TYPE: PRT | | |
| ; ORGANISM: Pseudomonas aeruginosa | | |
| US-09-252-991A-25685 | | |
| Query Match | 18.0%; Score 59.5; DB 4; Length 688; | |
| Best Local Similarity | 32.3%; Pred. No. 8.4; | |
| Matches | 21; Conservative 6; Mismatches 17; Indels 21; Gaps 3; | |
| Qy | 2 GIISRLMSVVEELKRDHAEMQAGCGLOQTEDHLMRRS 40 | |
| Db | 299 GADQRFVLVDEQDRLSRAVAGVGDHPLEALLELALHAGAGLQQAQAEVHPQDHLQRRHF 358 | |
| | | |
| Qy | 41 ASIDA 45 | |
| | | |

GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540.236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2236
; LENGTH: 260
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-2236

Query Match 17.4%; Score 57.5; DB 4; Length 260;
Best Local Similarity 37.3%; Pred. No. 4.5;
Matches 19; Conservative 7; Mismatches 14; Indels 11; Gaps 3;

Qy 13 ELKP-DHAEMQAGCGCLOTHDHLMPRRSAFASLDV--NARLMSALTPAXEY 60
Db 81 DLERADFAQMKLAGLIEK-----FGHLDGVLHNAAILGALTPLEMY 123

RESULT 12
US-09-126-420A-20
; Sequence 20, Application US/09126420A
; Patent No. 6376753
; GENERAL INFORMATION:
; APPLICANT: BATAARD, YANNICK
; APPLICANT: ROBINEAU, TIBURCE
; APPLICANT: DURST, FRANCIS
; APPLICANT: WERCK-REICHAUT, DANIELE
; APPLICANT: DIDIERJEAN, LUC
; TITLE OF INVENTION: PURIFIED CYTOCHROME P450 CYP78B1 FROM HELIANTHUS
; TITLE OF INVENTION: TUBEROSUS AND ITS APPLICATIONS AS BIOCATALYST IN
; TITLE OF INVENTION: PARTICULAR FOR THE DEGRADATION OF ENVIRONMENTAL
; TITLE OF INVENTION: POLLUTANTS AND FOR ALTERING THE RESISTANCE OF PLANTS S
; FILE REFERENCE: 03715.0032
; CURRENT APPLICATION NUMBER: US/09/126.420A
; CURRENT FILING DATE: 1998-07-30
; PRIOR FILING DATE: 60/054.351
; PRIOR FILING DATE: 1997-07-31
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 20
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Persea americana
US-09-126-420A-20

Query Match 17.3%; Score 57; DB 3; Length 471;
Best Local Similarity 34.4%; Pred. No. 12;
Matches 11; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

Qy 7 LMSVEBELKRDHAEQMAGCGCLOTHDHLMPRRS 38
Db 230 LTGMCDARKRNHGBLDAFVHDVIDDHLRSKA 261

RESULT 13
US-09-538-092-1285
; Sequence 1285, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538.092
; CURRENT FILING DATE: 2000-03-29
; PRIOR FILING DATE: 60/127.352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965

; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSeqFormatter Version 0.9
; SEQ ID NO 1285
; LENGTH: 620
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number Q08379
US-09-538-092-1285

Query Match 17.3%; Score 57; DB 4; Length 620;
Best Local Similarity 35.7%; Pred. No. 17;
Matches 15; Conservative 5; Mismatches 10; Indels 12; Gaps 1;

Qy 4 ISRLMSVEBELKRDHAEQMAG-----CGLQTHDHL 33
Db 153 ISRLSQNELKELQELAEQSGFVKLTWENMEITSLQSEQHV 194

RESULT 14
US-09-198-452A-67
; Sequence 67, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffeais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198.452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 67
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-67

Query Match 17.1%; Score 56.5; DB 4; Length 526;
Best Local Similarity 28.3%; Pred. No. 16;
Matches 15; Conservative 12; Mismatches 19; Indels 7; Gaps 2;

Qy 6 RLMSVEBELKRDHAEQMAGCGCLOTE-----DHLMPRRSAFAS--LDVAVNARLM 51
Db 365 RFQSLNQDLQNVQECQKATGSEVSAIRDLREQITEFETQGLDVIKEELL 417

RESULT 15
US-09-499-302A-7
; Sequence 7, Application US/09499302A
; Patent No. 6369212
; GENERAL INFORMATION:
; APPLICANT: BOUNG-JUN, OH
; APPLICANT: MOON, KYUNG KO
; APPLICANT: YOUNG, SOON KIM
; TITLE OF INVENTION: A CYTOCHROME P450 GENE HIGHLY EXPRESSED IN THE
; TITLE OF INVENTION: INCOMPATIBLE INTERACTION
; FILE REFERENCE: 10324/P64443USO
; CURRENT APPLICATION NUMBER: US/09/499.302A
; CURRENT FILING DATE: 2000-02-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 7
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Persea sp.
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (1)
; OTHER INFORMATION: variable or unknown amino acid

```

; NAME/KEY: MOD_RES
; LOCATION: (22)
; OTHER INFORMATION: variable or unknown amino acid
; NAME/KEY: MOD_RES
; LOCATION: (30)
; OTHER INFORMATION: variable or unknown amino acid
; NAME/KEY: MOD_RES
; LOCATION: (49)
; OTHER INFORMATION: variable or unknown amino acid
; NAME/KEY: MOD_RES
; LOCATION: (193)
; OTHER INFORMATION: variable or unknown amino acid
; NAME/KEY: MOD_RES
; LOCATION: (226)
; OTHER INFORMATION: variable or unknown amino acid
; NAME/KEY: MOD_RES
; LOCATION: (240)
; OTHER INFORMATION: variable or unknown amino acid
; NAME/KEY: MOD_RES
; LOCATION: (247)
; OTHER INFORMATION: variable or unknown amino acid
; NAME/KEY: MOD_RES
; LOCATION: (262)
; OTHER INFORMATION: variable or unknown amino acid
; NAME/KEY: MOD_RES
; LOCATION: (278)
; OTHER INFORMATION: variable or unknown amino acid
; US-09-499-302A-7

```

```

Query Match      17.0%  Score 56;  DB 3;  Length 502;
Best Local Similarity 34.4%  Pred.No. 18;
Matches 11;  Conservative 7;  Mismatches 14;  Indels 0;  Gaps 0;

```

```

Qy      7  LMSVEELKRDHAEMQAGCGLQTEDHLMPPRS 38
Db      230  LTGMDARLKRKHGELDAXVDHVLDDHLISRKA 261

```

```

Search completed: November 10, 2004, 12:32:32
Job time : 20.9552 secs

```

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 15:53:52 ; Search time 64.6219 Seconds
(without alignments)
366.225 Million cell updates/sec

Title: US-10-092-750-20

Perfect score: 330

Sequence: 1 KGIISRLMSVEELKRDHAE.....ARLMSALTPAXRYVXHCXPL 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/prodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep.*
- 16: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep.*
- 17: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/prodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 324 | 98.2 | 67 | 14 | US-10-092-750-20 |
| 2 | 154.5 | 46.8 | 1054 | 15 | Sequence 108, Appl |
| 3 | 66.5 | 20.2 | 964 | 16 | Sequence 2, Appl |
| 4 | 64.5 | 19.5 | 621 | 17 | Sequence 240989, |
| 5 | 63 | 19.1 | 183 | 16 | Sequence 172345, |
| 6 | 62.5 | 18.9 | 203 | 16 | Sequence 191237, |
| 7 | 62 | 18.8 | 530 | 14 | Sequence 6, Appl |
| 8 | 62 | 18.8 | 1016 | 14 | US-10-371-905A-6 |
| 9 | 61.5 | 18.6 | 395 | 14 | US-10-371-905A-4 |
| 10 | 61.5 | 18.6 | 401 | 16 | Sequence 17471, A |
| 11 | 61.5 | 18.6 | 406 | 10 | Sequence 1369, Ap |
| 12 | 61.5 | 18.6 | 406 | 10 | US-10-408-765A-1369 |
| 13 | 61.5 | 18.6 | 406 | 10 | US-09-782-816A-53 |
| 14 | 61.5 | 18.6 | 406 | 10 | US-09-925-298-676 |
| 15 | 61.5 | 18.6 | 406 | 10 | US-10-102-806-676 |

| | | | | | | |
|----|------|------|-------|----|----------------------|-------------------|
| 14 | 61.5 | 18.6 | 2446 | 13 | US-10-087-192-2028 | Sequence 2028, Ap |
| 15 | 61.5 | 18.6 | 2920 | 17 | US-10-038-854-71 | Sequence 71, Appl |
| 16 | 60 | 18.2 | 128 | 17 | US-10-425-115-325553 | Sequence 325553, |
| 17 | 59.5 | 18.0 | 191 | 14 | US-10-189-346-52 | Sequence 52, Appl |
| 18 | 59.5 | 18.0 | 208 | 14 | US-10-489-346-53 | Sequence 53, Appl |
| 19 | 59.5 | 18.0 | 412 | 15 | US-10-282-122A-68320 | Sequence 68320, A |
| 20 | 59 | 17.9 | 403 | 16 | US-10-437-963-160144 | Sequence 160144, |
| 21 | 58.5 | 17.7 | 715 | 15 | US-10-282-122A-78494 | Sequence 78494, A |
| 22 | 58.5 | 17.7 | 984 | 16 | US-10-437-963-105777 | Sequence 105777, |
| 23 | 58.5 | 17.7 | 1250 | 16 | US-10-437-963-188407 | Sequence 188407, |
| 24 | 58.5 | 17.7 | 1827 | 14 | US-10-369-493-5368 | Sequence 5368, Ap |
| 25 | 58 | 17.6 | 125 | 15 | US-10-425-114-57602 | Sequence 57602, A |
| 26 | 58 | 17.6 | 216 | 15 | US-10-424-599-240457 | Sequence 240457, |
| 27 | 58 | 17.6 | 360 | 10 | US-09-372-348-9 | Sequence 9, Appl |
| 28 | 58 | 17.6 | 391 | 16 | US-10-437-963-168806 | Sequence 168806, |
| 29 | 58 | 17.6 | 530 | 10 | US-09-372-348-11 | Sequence 11, Appl |
| 30 | 58 | 17.6 | 530 | 14 | US-10-107-6958-2 | Sequence 2, Appl |
| 31 | 58 | 17.6 | 530 | 14 | US-10-371-905A-10 | Sequence 10, Appl |
| 32 | 58 | 17.6 | 530 | 14 | US-10-371-905A-12 | Sequence 12, Appl |
| 33 | 58 | 17.6 | 530 | 14 | US-10-371-905A-16 | Sequence 16, Appl |
| 34 | 58 | 17.6 | 530 | 14 | US-10-371-905A-18 | Sequence 18, Appl |
| 35 | 58 | 17.6 | 565 | 14 | US-10-371-905A-20 | Sequence 20, Appl |
| 36 | 58 | 17.6 | 574 | 14 | US-10-371-905A-14 | Sequence 14, Appl |
| 37 | 58 | 17.6 | 633 | 14 | US-10-371-905A-24 | Sequence 24, Appl |
| 38 | 58 | 17.6 | 715 | 14 | US-10-371-905A-22 | Sequence 22, Appl |
| 39 | 58 | 17.6 | 1021 | 14 | US-10-371-905A-8 | Sequence 8, Appl |
| 40 | 58 | 17.6 | 19652 | 15 | US-10-084-846A-7 | Sequence 7, Appl |
| 41 | 57.5 | 17.4 | 80 | 15 | US-10-424-599-252794 | Sequence 252794, |
| 42 | 57.5 | 17.4 | 376 | 14 | US-10-156-761-11851 | Sequence 11851, A |
| 43 | 57.5 | 17.4 | 410 | 15 | US-10-282-122A-69704 | Sequence 69704, A |
| 44 | 57.5 | 17.4 | 419 | 15 | US-10-424-599-160165 | Sequence 160165, |
| 45 | 57.5 | 17.4 | 862 | 14 | US-10-369-493-19762 | Sequence 19762, A |

ALIGNMENTS

RESULT 1

US-10-092-750-20
; Sequence 20, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-XL
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092.750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 58, 62, 65
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-092-750-20

Query Match 98.2% Score 324; DB 14; Length 67;

Best Local Similarity 100.0%; Pred. No. 3.8e-35; Mismatches 0; Indels 0; Gaps 0;

Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KGIISRLMSVEELKRDHAEVQACGLQTEHLMPSAFASLDAVNARLMSALTPAXRY 60

Db 1 KGIISRLMSVEELKRDHAEVQACGLQTEHLMPSAFASLDAVNARLMSALTPAXRY 60

Qy 61 VVHCXPL 67

|||||

Db 61 VHXKPL 67

RESULT 2

US-10-001-885-108

Sequence 108, Application US/10001885

Publication No. US20040059319A1

GENERAL INFORMATION:

APPLICANT: Salceda, Susana

APPLICANT: Macina, Roberto

APPLICANT: Recipon, Herve

APPLICANT: Cafferey, Robert

APPLICANT: Sun, Yongming

APPLICANT: Liu, Chenghua

TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro

FILE REFERENCE: DEX-0279

CURRENT APPLICATION NUMBER: US/10/001,885

CURRENT FILING DATE: 2001-11-20

PRIOR APPLICATION NUMBER: 60/252,061

PRIOR FILING DATE: 2000-11-20

PRIOR APPLICATION NUMBER: 60/253,257

PRIOR FILING DATE: 2000-11-27

NUMBER OF SEQ ID NOS: 167

SOFTWARE: PatentIn version 3.1

SEQ ID NO 108

LENGTH: 1054

TYPE: PRT

ORGANISM: Homo sapien

US-10-001-885-108

Query Match 46.8%; Score 154.5; DB 15; Length 1054;

Best Local Similarity 52.7%; Pred No. 3.3e-11;

Matches 39; Conservative 4; Mismatches 12; Indels 19; Gaps 2;

QY 1 KGIISRLMSVEELKRDHAENQAG-----CGIQTEDH-----LMPRSFA 41

Db 976 KGIISRLMSVEELKRDHAENQAG-----CGIQTEDH-----LMPRSFA 41

QY 42 SLDAVNARLMSALT 55

Db 1036 SLDAVNARLMSALT 1049

RESULT 3

US-10-239-734-2

Sequence 2, Application US/10239734

Publication No. US20040161746A1

GENERAL INFORMATION:

APPLICANT: GENOX RESEARCH, INC.

APPLICANT: JAPAN AS REPRESENTED BY GENERAL DIRECTOR OF AGENCY OF NATIONAL CENTER FOR

APPLICANT: CHILD HEALTH AND DEVELOPMENT

APPLICANT: Matsumoto, Yoshiko

APPLICANT: Tsujimoto, Gozoh

APPLICANT: Nagasu, Takeshi

APPLICANT: Sugita, Yuji

APPLICANT: Oshida, Tadashi

APPLICANT: Imai, Yukiko

TITLE OF INVENTION: Method of Testing For Allergic Disease

FILE REFERENCE: SHIMIZU-07379

CURRENT APPLICATION NUMBER: US/10/239,734

CURRENT FILING DATE: 2002-09-24

PRIOR APPLICATION NUMBER: PCT/JP01/11286

PRIOR FILING DATE: 2001-12-21

PRIOR APPLICATION NUMBER: 2000-389476 JP

PRIOR FILING DATE: 2000-12-21

NUMBER OF SEQ ID NOS: 26

SOFTWARE: PatentIn version 3.1

SEQ ID NO 2

LENGTH: 964

TYPE: PRT

ORGANISM: Homo sapiens

US-10-239-734-2

Query Match 20.2%; Score 66.5; DB 16; Length 964;

Best Local Similarity 36.7%; Pred No. 14;

Matches 18; Conservative 10; Mismatches 18; Indels 3; Gaps 1;

QY 10 VEEELKRDHAENQAGCG---LQTEDHLMPPRSFAFASLDVAVNARLMSALT 55

Db 563 VPEELKRDYVDEQTDGDPKVSRETLKSRKSDYSLNKVNAPILTNTT 611

RESULT 4

US-10-425-115-240989

Sequence 240989, Application US/10425115

Publication No. US20040214272A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants

FILE REFERENCE: 38-21(53222)B

CURRENT APPLICATION NUMBER: US/10/425,115

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 369326

SEQ ID NO 240989

LENGTH: 621

TYPE: PRT

ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Clone ID: MET4577_151360C.1.pep

US-10-425-115-240989

Query Match 19.5%; Score 64.5; DB 17; Length 621;

Best Local Similarity 30.6%; Pred No. 15;

Matches 22; Conservative 11; Mismatches 24; Indels 15; Gaps 3;

QY 7 LMSVEEEL-----KRDHAENQAGCGIQTEDHLMPPRSFAFASLDVAVNARLM-----S 52

Db 42 ILSLEESIGKFEVHYMQKRD-ATLKEGWKLQREKEAMLKAMHDSLRSKAELLTKFSRS 100

QY 53 ALTPAXRYVXHC 64

Db 101 ADVDPSTYVSHC 112

RESULT 5

US-10-437-963-172345

Sequence 172345, Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov, Andrey A.

APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53221)B

CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 172345

LENGTH: 183

TYPE: PRT

ORGANISM: Oryza sativa

FEATURE:

NAME/KEY: unsure

LOCATION: (1)..(183)

OTHER INFORMATION: unsure at all Xaa locations

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT4530_70490C.1.pep
US-10-437-963-172345

Query Match 19.1%; Score 63; DB 16; Length 183;
Best Local Similarity 37.5%; Pred. No. 5.3;
Matches 18; Conservative 5; Mismatches 13; Indels 12; Gaps 1;

QY 6 RLMSVEELKRDHAEMQAGCGL-----QTEDHLMPPRSFAFA 41
DB 41 RAAELEERLQVAESQAWGLARSNEAAGVATLIDHLLRAAAA 88

RESULT 6

US-10-437-963-191237
Sequence 191237, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 191237
LENGTH: 203
TYPE: PRT
ORGANISM: Oryza sativa
NAME/KEY: unsure
LOCATION: (1), (203)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_87574C.1.pep
US-10-437-963-191237

Query Match 18.9%; Score 62.5; DB 16; Length 203;
Best Local Similarity 32.7%; Pred. No. 7;
Matches 17; Conservative 11; Mismatches 23; Indels 1; Gaps 1;

QY 6 RLMSVEELKRDHAEMQAGCGLQTEDHLMPPRSFAFASLDVAVNRLMSALTPA 57
DB 142 RAAELEERLQVAESQAWGLAXQNEAV-AAGLRATLDHLLRAVAPAQA 192

RESULT 7

US-10-371-905A-6
Sequence 6, Application US/10371905A
Publication No. US20030224969A1
GENERAL INFORMATION:
APPLICANT: Hahn, Chang
APPLICANT: Liu, Hong
TITLE OF INVENTION: Human Analogs of Murine Deubiquitinating Protease Genes
FILE REFERENCE: USAV2001/0185
CURRENT APPLICATION NUMBER: US/10/371,905A
CURRENT FILING DATE: 2003-02-20
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 530
TYPE: PRT
ORGANISM: Homo sapiens
NAME/KEY: MISC FEATURE
OTHER INFORMATION: human DUB4.1b
US-10-371-905A-6

Query Match 18.8%; Score 62; DB 14; Length 530;
Best Local Similarity 26.2%; Pred. No. 26;
Matches 22; Conservative 12; Mismatches 26; Indels 24; Gaps 3;

QY 1 KGIISRLMSVEE-----ELKRDHAEMQAGCGLQTEDHLMPPRSFAFASLD----- 44
DB 387 RGREPRALGVEDTDRRTATQCGELKRDHPCLQAP---ELDEHLVERATQESTLIDHWKFLQEQ 443
QY 45 -----AVNARLMSALTTPAXRVVXH 63
DB 444 NTKPEFNVRRVEGTVPDPVLVIH 467

RESULT 8

US-10-371-905A-4
Sequence 4, Application US/10371905A
Publication No. US20030224969A1
GENERAL INFORMATION:
APPLICANT: Hahn, Chang
APPLICANT: Liu, Hong
TITLE OF INVENTION: Human Analogs of Murine Deubiquitinating Protease Genes
FILE REFERENCE: USAV2001/0185
CURRENT APPLICATION NUMBER: US/10/371,905A
CURRENT FILING DATE: 2003-02-20
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 1016
TYPE: PRT
ORGANISM: Homo sapiens
NAME/KEY: MISC FEATURE
OTHER INFORMATION: human DUB4.1a
US-10-371-905A-4

Query Match 18.8%; Score 62; DB 14; Length 1016;
Best Local Similarity 26.2%; Pred. No. 58;
Matches 22; Conservative 12; Mismatches 26; Indels 24; Gaps 3;

QY 1 KGIISRLMSVEE-----ELKRDHAEMQAGCGLQTEDHLMPPRSFAFASLD----- 44
DB 873 RGREPRALGVEDTDRRTATQCGELKRDHPCLQAP---ELDEHLVERATQESTLIDHWKFLQEQ 929
QY 45 -----AVNARLMSALTTPAXRVVXH 63
DB 930 NTKPEFNVRRVEGTVPDPVLVIH 953

RESULT 9

US-10-369-493-17471
Sequence 17471, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xiaofeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 17471
LENGTH: 395
TYPE: PRT
ORGANISM: Bacillus halodurans
US-10-369-493-17471

Query Match 18.6%; Score 61.5; DB 14; Length 395;
Best Local Similarity 26.2%; Pred. No. 21;
Matches 22; Conservative 10; Mismatches 27; Indels 25; Gaps 2;
QY 2 GIIISRLMSVEEL---KRDHAEAGGCGGCTEDHLMPPRS----- 38
DB 116 GIIISIVRGDAVFSKLNHSAIVDGCQGRADHLRFHMDHLETLQKSPHKOKLIV 175
QY 39 --AFASLDVAVNARMSALTTPAXRY 60
DB 176 VDALFSMDGDHANLHDLVTLKERY 199

RESULT 10
US-10-408-765A-1369
; Sequence 1369, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Wainock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1369
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1369

Query Match 18.6%; Score 61.5; DB 16; Length 401;
Best Local Similarity 31.3%; Pred. No. 22;
Matches 21; Conservative 8; Mismatches 29; Indels 9; Gaps 2;
QY 6 RLMSVEEELKRDH-----AEMQAGCGGCTEDHLMPPRSA--FASLDVAVNARMSALTTP 56
DB 230 RLTELETAVRCDDQDAQNPLSAGLOGACIMETVELLOAKVSALDLAVLDQVEARLQSVLKG 289
QY 57 AXRYVXH 63
DB 290 VNEIAKH 296

RESULT 11
US-09-782-816A-53
; Sequence 53, Application US/09782816A
; Publication No. US20030032771A1
; GENERAL INFORMATION:
; APPLICANT: Sharp, David J.
; APPLICANT: Rogers, Gregory C.
; APPLICANT: Scholey, Jonathan M.
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF CELLULAR
; TITLE OF INVENTION: PROLIFERATION
; FILE REFERENCE: U069.001A
; CURRENT APPLICATION NUMBER: US/09/782.816A
; CURRENT FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-816A-53

Query Match 18.6%; Score 61.5; DB 10; Length 406;

Best Local Similarity 31.3%; Pred. No. 22;
Matches 21; Conservative 8; Mismatches 29; Indels 9; Gaps 2;
QY 6 RLMSVEEELKRDH-----AEMQAGCGGCTEDHLMPPRSA--FASLDVAVNARMSALTTP 56
DB 235 RLTELETAVRCDDQDAQNPLSAGLOGACIMETVELLOAKVSALDLAVLDQVEARLQSVLKG 294
QY 57 AXRYVXH 63
DB 295 VNEIAKH 301

RESULT 12
US-09-925-298-676
; Sequence 676, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 676
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (5)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (6)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (16)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-298-676

Query Match 18.6%; Score 61.5; DB 9; Length 465;
Best Local Similarity 31.3%; Pred. No. 26;
Matches 21; Conservative 8; Mismatches 29; Indels 9; Gaps 2;

QY 6 RLMSVEEELKRDH-----AEMQAGCGGCTEDHLMPPRSA--FASLDVAVNARMSALTTP 56
DB 294 RLTELETAVRCDDQDAQNPLSAGLOGACIMETVELLOAKVSALDLAVLDQVEARLQSVLKG 353
QY 57 AXRYVXH 63
DB 354 VNEIAKH 360

RESULT 13
US-10-102-806-676
; Sequence 676, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12


```

; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 676
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (5)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (6)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (116)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-10-102-806-676

Query Match      18.6%; Score 61.5; DB 14; Length 465;
Best Local Similarity 31.3%; Pred. No. 26;
Matches 21; Conservative 8; Mismatches 29; Indels 9; Gaps 2

QY    6 RLMSVEELKRDH-----AEMAGCGLQTEDHLMPPRSA--FASLDVNRRLMSALTP 56
Db    294 RLTLETAVRCDDQAQNPLSAGLCIMETVELLOAKVSALDLAVLDQVEARLQSVLGK 353
QY    57 AXRYVKX 63
Db    354 VNEIAKH 360

RESULT 14
US-10-087-192-2028
; Sequence 2028, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FascSeq for Windows Version 4.0
; SEQ ID NO 2028
; LENGTH: 2446
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-087-192-2028

Query Match      18.6%; Score 61.5; DB 13; Length 2446;
Best Local Similarity 25.8%; Pred. No. 2e+02;
Matches 17; Conservative 17; Mismatches 23; Indels 9; Gaps 2

QY    3 IISRLMSYE---EELKRDBAEWQ-----AGCGLQTEDHLMPPRSASFASILDVNRRLMSA 53
Db    1469 IVEELSAVELTNSDIKKOLSRPQPQLVRQCACSEPKDGLSGSSSFSLSPSSSQDPFS 1528
QY    54 LTPAXR 59
Db    1529 VSPSSR 1534

RESULT 15
US-10-038-854-71
; Sequence 71, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
```

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:36:51 ; Search time 13.5448 Seconds
(without alignments)
475.942 Million cell updates/sec

Title: US-10-092-750-20

Perfect score: 330

Sequence: 1 KGIISRLMSVEELKRDHAE.....ARLMSALTPAXRYVXHCXPL 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 72 | 21.8 | 1249 | 2 T14270 | Ras-GTPase activat |
| 2 | 72 | 21.8 | 1293 | 2 T14259 | ras GTPase-activat |
| 3 | 63 | 19.1 | 2117 | 2 T36180 | CDA peptide synthe |
| 4 | 61.5 | 18.6 | 335 | 2 C84138 | 8-amino-7-oxononan |
| 5 | 61.5 | 18.6 | 2500 | 1 WMRUE2 | HIV-EP2 enhancer-B |
| 6 | 60.5 | 18.3 | 185 | 2 F84182 | hypothetical prote |
| 7 | 60 | 18.2 | 456 | 2 F81592 | chromosomal replic |
| 8 | 59.5 | 18.0 | 1336 | 2 T23310 | hypothetical prote |
| 9 | 59 | 17.9 | 708 | 2 H82751 | ATP-dependent heli |
| 10 | 59 | 17.9 | 968 | 2 T51523 | clpB heat shock pr |
| 11 | 59 | 17.9 | 1387 | 2 T16511 | hypothetical prote |
| 12 | 58.5 | 17.7 | 644 | 2 T46211 | hypothetical prote |
| 13 | 58.5 | 17.7 | 715 | 2 AH0042 | formate dehydrogen |
| 14 | 58.5 | 17.7 | 1827 | 2 T16270 | hypothetical prote |
| 15 | 58 | 17.6 | 354 | 2 T22934 | hypothetical prote |
| 16 | 58 | 17.6 | 418 | 2 T34660 | probable sorbitol |
| 17 | 57.5 | 17.4 | 376 | 2 B41870 | DNA-directed DNA p |
| 18 | 57.5 | 17.4 | 536 | 1 JQBST7 | head-to-tail joini |
| 19 | 57 | 17.3 | 177 | 2 D83801 | sporulation initia |
| 20 | 57 | 17.3 | 358 | 1 QQBET1 | hypothetical prote |
| 21 | 57 | 17.3 | 471 | 1 A35867 | cytochrome P450 71 |
| 22 | 57 | 17.3 | 502 | 2 T25256 | cytochrome P-450IX |
| 23 | 57 | 17.3 | 620 | 2 JH0821 | 95K golgi antigen |
| 24 | 57 | 17.3 | 876 | 2 S20531 | hypothetical prote |
| 25 | 56.5 | 17.1 | 735 | 2 G86496 | hypothetical prote |
| 26 | 56.5 | 17.1 | 735 | 2 B81546 | hypothetical prote |
| 27 | 56.5 | 17.1 | 1551 | 2 AB2410 | WD-repeat protein |
| 28 | 56.5 | 17.1 | 1902 | 2 S06997 | lactocepin (EC 3.4 |
| 29 | 56.5 | 17.1 | 2442 | 2 T08621 | centrosome associa |

30 56 17.0 252 2 AH0932 thiamin biosynthes
31 56 17.0 321 2 T41036 ribose-phosphate p
32 56 17.0 369 2 T21556 hypothetical prote
33 56 17.0 3328 2 T30835 breast cancer tumo
34 56 17.0 3329 2 T42205 breast cancer susc
35 56 17.0 3329 2 T30904 breast cancer tumo
36 55.5 16.8 375 2 AC0337 probable D-isomer
37 55.5 16.8 585 2 D90880 hypothetical prote
38 55.5 16.8 585 2 E85738 hypothetical prote
39 55.5 16.8 585 2 E64892 probable membrane
40 55.5 16.8 600 2 H95379 probable adenine d
41 55.5 16.8 762 1 A46590 protein kinase (EC
42 55.5 16.8 1792 2 T13939 myosin V - fruit f
43 55 16.7 214 2 E83431 type III export pr
44 55 16.7 533 2 A39092 glucuronosyltransf
45 55 16.7 558 2 F75382 probable propionyl

ALIGNMENTS

RESULT 1

T14270

Ras-GTPase activating protein Syngap-b - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T14270

R:Kim, J.H.; Liao, D.; Lau, L.F.; Haganir, R.L.

Neuron 20, 683-691, 1998

A:Title: Syngap: a synaptic RasGAP that associates with the PSD-95/SAP90 protein family.

A:Reference number: Z17950; MUID:98240917; PMID:9581761

A:Accession: T14270

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1249 <KIM>

A:Cross-references: EMBL:AF058790; NID:G3722228; PID:G3722229; PIDN:AAC63511.1

Query Match

Best Local Similarity 71.8%; Score 72; DB 2; Length 1249;

Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 KGIISRLMSVEELKRDHAE 21

DB 1201 KSIIGRLMLVEELRRDHPAM 1221

RESULT 2

T14259

ras GTPase-activating protein, synaptic - rat

N:Alternate names: protein Syngap

C:Species: Rattus norvegicus (Norway rat)

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C:Accession: T14259

R:Chen, H.J.; Rojas-Soto, M.; Kennedy, M.B.

Submitted to the EMBL Data Library, February 1998

A:Description: A synaptic Ras GTPase-activating protein (p135 Syngap) inhibited by Cam k

A:Reference number: Z17942

A:Accession: T14259

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1293 <CHE>

A:Cross-references: UNIPROT:Q9QUH6; EMBL:AF048976; NID:G2935447; PID:G2935448; PIDN:AAC08

A:Experimental source: strain Sprague-Dawley

Query Match

Best Local Similarity 71.4%; Score 72; DB 2; Length 1293;

Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 KGIISRLMSVEELKRDHAE 21

DB 1245 KSIIGRLMLVEELRRDHPAM 1265

```
RESULT 3
T36180
CDA peptide synthetase III - Streptomyces coelicolor (fragment)
C:Species: Streptomyces coelicolor
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T36180
R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrall, B.G.; Rajandream, M.
submitted to the EMBL Data Library, March 1999
A:Reference number: 221600
A:Accession: T36180
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2117 <SAU>
A:Cross-references: UNIPROT:Q8GUX2; EMBL:AL035707; PIDN:CAB38876.1; GSPDB:GN00070; SCOPED
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: cdaPS3; SCOPED:SCF63.01
C:Keywords: carrier protein; phosphopantetheine; phosphoprotein
F:208-662/Domain: acetate-CoA ligase homology <ACLI>
F:678-746/Domain: acyl carrier protein homology <ACP1>
F:1282-1725/Domain: acetate-CoA ligase homology <ACL2>
F:1740-1808/Domain: acyl carrier protein homology <ACP2>
F:710,1772/Binding site: phosphopantetheine (Ser) (covalent) #status predicted
Query Match 19.1%; Score 63; DB 2; Length 2117;
Best Local Similarity 39.6%; Pred. No. 35;
Matches 19; Conservative 5; Mismatches 18; Indels 6; Gaps 2;
QY 17 DHAEMQAGCGLTEDHLMPPRSAPASLDV---NARLMSALTPAXRY 60
DB 622 DEAAARALALSJDFPMWP--SAFVALDALPLSPNGKLDRAALPAPTY 667

RESULT 4
C84138
8-amino-7-oxononanoate synthase bioF [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: C84138
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-395 <STO>
A:Cross-references: UNIPROT:Q9K625; GB:AP001520; GB:BA000004; NID:G10176401; PIDN:BAB076
A:Experimental source: strain C-125
C:Genetics:
A:Gene: bioF
C:Superfamily: 5-aminolevulinic synthase; glycine C-acetyltransferase homology
C:Keywords: phosphoprotein; pyridoxal phosphate
F:240/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted
Query Match 18.6%; Score 61.5; DB 2; Length 395;
Best Local Similarity 26.2%; Pred. No. 8.1;
Matches 22; Conservative 10; Mismatches 27; Indels 25; Gaps 2;
QY 2 GIISRLMSVSEEL---KRDHAEQAGCGLTEDHLMPPRS----- 38
DB 116 GIISVIGRGDGVFSDKLNHASIVDGCQLSRADLRFRHNDMDHLETLQKSPHKQKLV 175

QY 39 --APASLDVARNALMSALTPAXRY 60
DB 176 VDALFSDGDDGHANLHDLVTLKERY 199

RESULT 5
WHU052
HIV-Ep2 enhancer-binding protein - human
N:Alternate names: finger protein, 275K; human immunodeficiency virus enhancer-binding p
C:Species: Homo sapiens (man)
C>Date: 31-Mar-1993 #sequence_revision 07-Jul-1995 #text_change 09-Jul-2004
C:Accession: S26661; A39829; A38253
R:van't Veer, L.J.; Lutz, P.M.; Isselbacher, K.J.; Bernards, R.
Proc. Natl. Acad. Sci. U.S.A. 89, 8971-8975, 1992
A:Title: Structure and expression of major histocompatibility complex-binding protein 2,
A:Reference number: S26661; MUID:93028387; PMID:11409593
A:Accession: S26661
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-2500 <VAN>
A:Cross-references: UNIPROT:Q02646; EMBL:X65644; NID:G38259; PIDN:CAA45596.1; PID:G38260
R:Nomura, N.; Zhao, M.J.; Nagase, T.; Maekawa, T.; Iizuka, R.; Tabata, S.; Ishii, S.
J. Biol. Chem. 266, 8590-8594, 1991
A:Title: HIV-Ep2, a new member of the gene family encoding the human immunodeficiency vi
A:Reference number: A39829; MUID:91217105; PMID:2022670
A:Accession: A39829
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 568-2144, R', 2146-2500 <NOM>
A:Cross-references: GB:M60119; NID:G2661140; PIDN:AAB88218.1; PID:G182120
R:Rustgi, A.K.; Van't Veer, L.J.; Bernards, R.
Proc. Natl. Acad. Sci. U.S.A. 87, 8707-8710, 1990
A:Title: Two genes encode factors with NF-kappaB- and H2TF1-like DNA-binding properties.
A:Reference number: A38253; MUID:91062349; PMID:2247438
A:Accession: A38253
A:Molecule type: mRNA
A:Residues: 1851-1990 <RUS>
A:Cross-references: GB:M61744; GB:M33920; NID:G187404; PIDN:AAA36202.1; PID:G187405
C:Genetics:
A:Gene: GDB:HIVEP2
A:Cross-references: GDB:I29086; OMIM:143054
A:Map position: 623-624
C:Superfamily: HIV-Ep2 enhancer-binding protein
C:Keywords: DNA binding; duplication; metal binding; transcription regulation; zinc finger
F:242-298/Region: DNA binding #status predicted
F:245-265/Region: zinc finger CCHH motif
F:273-295/Region: zinc finger CCHH motif
F:991-997/Region: nuclear location signal
F:1004-1036/Region: serine-rich
F:1852-1908/Region: DNA binding #status predicted
F:1855-1875/Region: zinc finger CCHH motif
F:1883-1905/Region: zinc finger CCHH motif
F:1953-1977/Region: acidic
Query Match 18.6%; Score 61.5; DB 1; Length 2500;
Best Local Similarity 25.8%; Pred. No. 65;
Matches 17; Conservative 17; Mismatches 23; Indels 9; Gaps 2;
QY 3 IISRLMSVSE---EELKRDHAEQ-----AGCGLTEDHLMPPRSAPASLDVARNALMSA 53
DB 1593 IVEELSAVELTNSDKDLSPKQPLVROGCAEPKGLQSGSSFSLSLSPSSSQDYP 1582

QY 54 LTPAXR 59
DB 1583 VSPSSR 1588

RESULT 6
F84182
hypothetical protein Vng0219h [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: F84182
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leithausner, B.; Keller, K.; Cruz, R.; Hanson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: F84182
A>Status: preliminary
A:Molecule type: DNA
```

A;Residues: 1-185 <STO>
A;Cross-references: UNIPROT:Q9HS10; GB:AE004437; NID:g10579866; PIDN:AAG18826.1; GSPDB:G
C;Genetics:
A;Gene: VNG0219H

Query Match 18.3%; Score 60.5; DB 2; Length 185;
Best Local Similarity 32.8%; Pred. No. 4.6;
Matches 20; Conservative 10; Mismatches 30; Indels 1; Gaps 1;

QY 4 ISRLMSVEELKEDHAEMQAGCGLOTHLMPRSFAFASLDA-VNARLMSALTPAXRYX 62
DB 96 IETVADVDNEVYRWREMAHVDSNPEASTPEFAAARDAGMDARHVDALTAIFRAVR 155
QY 63 H 63
DB 156 Y 156

RESULT 7
F81692
Chromosomal replication initiator protein DnaA TC0521 [imported] - Chlamydia muridarum
C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: F81692
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: F81692
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-456 <TET>
A;Cross-references: UNIPROT:Q9PKB4; GB:AB002321; GB:AE002160; NID:g7190560; PIDN:AAF93936
A;Experimental source: strain Ni9g (MoPn)
C;Genetics:
A;Gene: TC0521
C;Superfamily: replication initiation protein dnaa

Query Match 18.2%; Score 60; DB 2; Length 456;
Best Local Similarity 33.3%; Pred. No. 15;
Matches 15; Conservative 9; Mismatches 17; Indels 4; Gaps 2;

QY 3 IISRLMSVEELKRDHAEMQAGCGLOTHLMPR-RSAPASLDAV 46
DB 413 VISSIRAIQKLEDDREDSVCGVQ---ELTKRLSSAYQSLDLI 454

RESULT 8
T23310
hypothetical protein K04D7.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T23310
R;Wild, A.
submitted to the EMBL Data Library, February 1996
A;Reference number: Z19725
A;Accession: T23310
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1336 <WIL>
A;Cross-references: UNIPROT:Q21216; EMBL:Z69664; PIDN:CAA93515.1; GSPDB:GN00022; CESP:K0
A;Experimental source: clone K04D7
C;Genetics:
A;Gene: CESP:K04D7.5
A;Map position: 4
A;Introns: 28/2; 116/3; 153/3; 270/1; 351/3; 579/3; 699/3; 1002/3; 1024/3; 1080/3; 1132/

Query Match 18.0%; Score 59.5; DB 2; Length 1336;
Best Local Similarity 27.0%; Pred. No. 57;
Matches 17; Conservative 9; Mismatches 18; Indels 19; Gaps 1;

QY 4 ISRLMSVEELKRDHAEMQAGCGLOTHLMPRSFAFASLD 44

DB 714 LSTWQLKEAEKRIADSKQSCGQTISALSNVSGDSNDVNGTHEKELMSTSIDSL 773
QY 45 AVN 47
DB 774 SIN 776

RESULT 9
H82751
ATP-dependent helicase XF0882 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: H82751
R;Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: H82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: H82751
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-708 <SIM>
A;Cross-references: UNIPROT:Q9PEZ6; GB:AE003927; GB:AE003849; NID:g9105783; PIDN:AAF8369;
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Al
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carreir, H.
as-Nato, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferrio, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohne
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.B.; Kuramae, E.E.; Laigre
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, E.C.; Miyaki, C.Y.;
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaku
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Zs
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF0882

Query Match 17.9%; Score 59; DB 2; Length 708;
Best Local Similarity 34.4%; Pred. No. 32;
Matches 21; Conservative 10; Mismatches 28; Indels 2; Gaps 1;

QY 1 KGIISRLMSVEELKRDHAEMQAGCGLOTHLMPRSFAFASLDVARNLMSALTPAXRY 60
DB 332 RGTOSRLAKPEVCEGFETLTALTSLHS--LLVPLRSAAAGLDACAARQEAALSRLRW 389
QY 61 V 61
DB 390 L 390

RESULT 10
T51523
clpB heat shock protein-like - Arabidopsis thaliana
N;Alternate names: protein T20K14.60
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C;Accession: T51523
R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000
A;Reference number: Z25394
A;Accession: T51523
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-968 <SAT>
A;Cross-references: UNIPROT:Q9LF37; EMBL:AL391143
A;Experimental source: cultivar Columbia; BAC clone T20K14
C;Genetics:
A;Map position: 5

A:Introns: 76/3; 149/3; 229/1; 294/2; 312/3; 363/1; 466/1; 862/3
A:Note: T20K14.60
C:Superfamily: endopeptidase Clp ATP-binding chain

Query Match 17.9%; Score 59; DB 2; Length 968;
Best Local Similarity 28.1%; Pred. No. 45;
Matches 18; Conservative 15; Mismatches 23; Indels 8; Gaps 2;

QY 1 KGIISRLMSVEELKRDHAEMQAGCGLOTHLMPRSA---FASLDVAVNARLMSALTPA 57

DB 542 RSVNSRLQSIKEEDRVNLEIQ-----QAREYDLNPAALKYGLNSLQRLNEAEKEL 596

QY 58 XRYV 61

DB 597 NEYL 600

RESULT 11

T16511

Hypothetical protein F59A6.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T16511

R:Nhan, M.

submitted to the EMBL Data Library, December 1995

A:Description: The sequence of C. elegans cosmid F59A6.

A:Accession: T16511

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1387 <NHA>

A:Cross-references: EMBL:U41994; NID:gl123047; PID:gl123054; PIDN:AAA83458.1; CESP:F59A6

C:Genetics:

A:Gene: CESP:F59A6.1

A:Introns: 18/3; 99/3; 195/3; 327/2; 367/1; 626/2; 931/2; 1191/3

Query Match

Best Local Similarity 17.9%; Score 59; DB 2; Length 1387;

Matches 19; Conservative 12; Mismatches 23; Indels 4; Gaps 2;

QY 3 IISRLMSVEELKRDHAEMQAGCGLOTHLMPRSAFASLDVAVNARLMSALTPAXRY 60

DB 1270 IESRPSREVRERDRKELRT---LOENBILIRLQVERE-LNAQLKSGITRANRF 1323

RESULT 12

T46211

Hypothetical protein T8P19.180 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004

C:Accession: T46211

R:Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Catolico, L.; Artiguenave, F.; Sa

submitted to the Protein Sequence Database, December 1999

A:Reference number: 223008

A:Accession: T46211

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-644 <CHO>

A:Cross-references: UNIPROT:Q9SMN2; EMBL:AL133315

A:Experimental source: cultivar Columbia; BAC clone T8P19

C:Genetics:

A:Map position: 3

A:Introns: 133/1; 373/3; 403/3; 496/3; 566/3

A:Note: T8P19.180

Query Match

Best Local Similarity 17.7%; Score 58.5; DB 2; Length 644;

Matches 15; Conservative 14; Mismatches 21; Indels 7; Gaps 2;

QY 3 IISRLMSVEELKRDHAEMQAGCGLOTHLMPRSAFASLDVAVNARLMS 52

DB 267 LVQNRLQVLVEKKDKMKIEELCSVKSEELNQLMEEEKKQKHRYELNAIQERTWS 323

RESULT 13

AH0042

formate dehydrogenase (EC 1.2.1.2) [imported] - Versinia pestis (strain CO92)

C:Species: Versinia pestis

C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

C:Accession: AH0042

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Hoiden, M.T.G.; Prentice, M.B.;

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, I.

Nature 413, 523-527, 2001

A:Title: Genome sequence of Versinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AH0042

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-715 <KUR>

A:Cross-references: UNIPROT:Q82IZ0; GB:AL590842; PIDN:CAC89203.1; PID:gl5978442; GSPDB:GN

C:Genetics:

A:Gene: fdhF

C:Superfamily: formate dehydrogenase

C:Keywords: oxidoreductase

Query Match

Best Local Similarity 17.7%; Score 58.5; DB 2; Length 715;

Matches 16; Conservative 5; Mismatches 16; Indels 9; Gaps 1;

QY 27 IQTEPHLMRPSASFASLDV-----NARLMSALTPAXRYVXH 63

DB 406 IQTEPDLAARAFADLVVVQDIFMTKTAELADVVLPATSWGHEH 451

RESULT 14

T16270

Hypothetical protein F35D11.11 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C:Accession: T16270

R:Fulton, B.

submitted to the EMBL Data Library, June 1995

A:Description: The sequence of C. elegans cosmid F35D11.

A:Reference number: Z18487

A:Accession: T16270

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1827 <FUL>

A:Cross-references: UNIPROT:Q20042; EMBL:U29381; NID:g868214; PID:g868224; PIDN:AAA68757.

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:F35D11.11

A:Introns: 76/2; 131/3; 159/3; 185/3; 221/3; 253/3; 320/1; 869/3; 1133/3; 1205/2; 1250/1;

Query Match

Best Local Similarity 17.7%; Score 58.5; DB 2; Length 1827;

Matches 18; Conservative 15; Mismatches 25; Indels 13; Gaps 2;

QY 4 ISRLMSVEELKRDHAEMQAGCGLOTHLMP-----RRSAFAS-----LDVAVNARL 50

DB 1612 VTAMKKFEEWKLEAEARAEKILRKELMQTAIEDLKRSFRSDTEKKEIGIRVRL 1671

QY 51 MSALTTPAXRYV 61

DB 1672 EREISALKRVH 1682

RESULT 15

T22934

Hypothetical protein F58G1.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T22934

R:Smve, R.

submitted to the EMBL Data Library, November 1996

Search completed: November 10, 2004, 12:29:25
Job time : 14.5448 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:27:01 ; Search time 73.3071 Seconds
(without alignments)
525.871 Million cell updates/sec

Title: US-10-092-750-20

Perfect score: 330
Sequence: 1 KGIIRLMSVBEELKRDHAE.....AFRLMSALTFAKYYVXHCXPL 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------------|---------------------|
| 1 | 153.5 | 46.5 | 475 | 2 Q6Y636 | Q6Y636 mus musculus |
| 2 | 153.5 | 46.5 | 475 | 2 AAP31233 | Aap31233 mus muscu |
| 3 | 153.5 | 46.5 | 811 | 2 O80T97 | O80T97 mus muscu |
| 4 | 153.5 | 46.5 | 967 | 2 Q8TDL2 | Q8tdl2 homo sapien |
| 5 | 153.5 | 46.5 | 996 | 2 Q6JTV1 | Q6jtv1 mus muscu |
| 6 | 153.5 | 46.5 | 996 | 2 Q6P730 | Q6p730 rattus norv |
| 7 | 153.5 | 46.5 | 996 | 2 Q924M9 | Q924m9 rattus norv |
| 8 | 153.5 | 46.5 | 996 | 2 AAH61865 | Aah61865 rattus no |
| 9 | 153.5 | 46.5 | 996 | 2 AAQ77379 | Aaq77379 mus muscu |
| 10 | 153.5 | 46.5 | 996 | 2 AAQ77380 | Aaq77380 mus muscu |
| 11 | 153.5 | 46.5 | 996 | 2 AAQ77381 | Aaq77381 mus muscu |
| 12 | 153.5 | 46.5 | 1036 | 2 Q9C0C0 | Q9c0c0 homo sapien |
| 13 | 153.5 | 46.5 | 1065 | 2 Q968E1 | Q968e1 homo sapien |
| 14 | 131.5 | 39.8 | 1139 | 1 NGAP HUMAN | Ngapfz2 homo sapien |
| 15 | 90.5 | 27.4 | 1556 | 1 GAP2 DROME | Q8t498 drosophila |
| 16 | 72 | 21.8 | 1293 | 1 SNGP RAT | Q9quh6 rattus norv |
| 17 | 72 | 21.8 | 1328 | 1 SNGP HUMAN | Q96pv0 homo sapien |
| 18 | 67 | 20.3 | 203 | 2 Q7R704 | Q7r704 plasmodium |
| 19 | 66.5 | 20.2 | 413 | 2 Q75N11 | Q75n11 homo sapien |
| 20 | 66.5 | 20.2 | 413 | 2 Q9NXR2 | Q9nxr2 homo sapien |
| 21 | 66.5 | 20.2 | 413 | 2 AAS02013 | Aas02013 homo sapi |
| 22 | 66.5 | 20.2 | 502 | 2 Q9H5C6 | Q9h5c6 homo sapien |
| 23 | 66.5 | 20.2 | 557 | 2 Q96JG6 | Q96jg6 homo sapien |
| 24 | 65 | 19.7 | 576 | 2 Q85K17 | Q85k17 chlamydomon |
| 25 | 63 | 19.1 | 2417 | 2 Q8CXJ2 | Q8cxj2 streptomyce |
| 26 | 62.5 | 18.9 | 504 | 2 Q8V503 | Q8v503 nocardioides |
| 27 | 62.5 | 18.9 | 352 | 2 Q9F829 | Q9f829 micromonospy |
| 28 | 62 | 18.8 | 495 | 2 Q8SUH1 | Q8suh1 encaphalito |
| 29 | 61.5 | 18.6 | 395 | 2 Q9K625 | Q9k625 bacillus ha |
| 30 | 61.5 | 18.6 | 400 | 1 DCT2 HUMAN | Q13561 homo sapien |
| 31 | 61.5 | 18.6 | 401 | 1 DCT2_MOUSE | Q99xj8 mus muscu |

ALIGNMENTS

RESULT 1

Q6Y636
ID Q6Y636 PRELIMINARY; PRT; 475 AA.
AC Q6Y636;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE DAB2-interacting protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=22760439; PubMed=12877983;
RA Homayouni R., Magdalen S., Keshvara L., Rice D.S., Curran T.;
RT "Interaction of Disabled-1 and the GTPase activating protein Dab2IP in mouse brain."
RL Brain Res. Mol. Brain Res. 115:121-129 (2003).
DR EMBL; AY178784; AAP31233.1;
DR GO; GO:0005515; F:protein binding; IPI.
FT NON_TER
SQ SEQUENCE 475 AA; 51159 MW; 5199721C6D7CF447 CRC64;
Query Match 46.5%; Score 153.5; DB 2; Length 475;
Best Local Similarity 65.5%; Pred. No. 3e-10;
Matches 36; Conservative 4; Mismatches 14; Indels 1; Gaps 1;

QY

1 KGIIRLMSVBEELKRDHAEVQAAGCGIQTEDHLMPPRSAFASLDANARLMSALT 55
415 KGIIRLMSVBEELKRDHAEVQAAGCGIQTEDHLMPPRSAFASLDANARLMSALT 468

Db

RESULT 2

AAP31233
ID AAP31233 PRELIMINARY; PRT; 475 AA.
AC AAP31233;
DT 02-MAR-2004 (Tremblrel. 27, Created)
DT 02-MAR-2004 (Tremblrel. 27, Last sequence update)
DT 02-MAR-2004 (Tremblrel. 27, Last annotation update)
DE DAB2-interacting protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=22760439; PubMed=12877983;
RA Homayouni R., Magdalen S., Keshvara L., Rice D.S., Curran T.;
RT "Interaction of Disabled-1 and the GTPase activating protein Dab2IP in mouse brain."
RL Brain Res. Mol. Brain Res. 115:121-129 (2003).

DR EMBL; AY178784; AAP31233.1; --
 FT NON TER 1
 SQ SEQUENCE 475 AA; 51159 MW; 5199721C6D7CF447 CRC64;
 Query Match 46.5%; Score 153.5; DB 2; Length 475;
 Best Local Similarity 65.5%; Pred. No. 3e-10;
 Matches 36; Conservative 4; Mismatches 14; Indels 1; Gaps 1;
 QY 1 KGIISRLMSVEEELKRDHAEMQACGCGLTEDHLMPPRSFASFLDANVARLMSALT 55
 Db 415 KGIISRLMSVEEELKRDHAEMQACGCGLTEDHLMPPRSFASFLDANVARLMSALT 468
 RESULT 3
 Q80T97
 ID Q80T97 PRELIMINARY; PRT; 811 AA.
 AC Q80T97;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE MKIAA1743 protein (Fragment).
 GN Name=Dab2ip; Synonyms=mKIAA1743;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22579291; PubMed=12693553;
 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
 RA Nakajima D., Nagase T., Ohara O., Koga H.;
 RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
 RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous
 RT cDNAs identified by screening of terminal sequences of cDNA clones
 RT randomly sampled from size-fractionated libraries.";
 RL DNA Res. 10:35-48(2003).
 DR EMBL; AK122548; BAC65830.1; --
 DR MGD; MGI:1916851; Dab2ip.
 DR GO; GO:0005515; F:protein binding; IPI.
 DR InterPro; IPR001936; RasGAP.
 DR Pfam; PF00616; RasGAP.
 DR SMART; SM00323; RasGAP; 1.
 DR PROSITE; PS50018; RAS_GTPASE_ACTIV_2; 1.
 FT NON_TER 1
 SQ SEQUENCE 811 AA; 89926 MW; 0DBE540BFD5B27A3 CRC64;
 Query Match 46.5%; Score 153.5; DB 2; Length 811;
 Best Local Similarity 65.5%; Pred. No. 5.4e-10;
 Matches 36; Conservative 4; Mismatches 14; Indels 1; Gaps 1;
 QY 1 KGIISRLMSVEEELKRDHAEMQACGCGLTEDHLMPPRSFASFLDANVARLMSALT 55
 Db 722 KGIISRLMSVEEELKRDHAEMQACGCGLTEDHLMPPRSFASFLDANVARLMSALT 775
 RESULT 4
 Q8TDL2
 ID Q8TDL2 PRELIMINARY; PRT; 967 AA.
 AC Q8TDL2;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE DOC-2/DAB2 interactive protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21945266; PubMed=11944990;
 RX Chen H., Pong R.C., Wang Z., Hsieh J.T.;

RT "Differential regulation of the human gene DAB2IP in normal and
 RT malignant prostatic epithelia: cloning and characterization.";
 RL Genomics 79:573-581(2002).
 DR EMBL; AF367051; AAM00371.1; --
 DR HSSP; P21359; INF1.
 DR InterPro; IPR000008; C2.
 DR InterPro; IPR008973; C2_CaLB.
 DR InterPro; IPR001936; RasGAP.
 DR InterPro; IPR008936; Rho_GAP.
 DR Pfam; PF00168; C2; 1.
 DR Pfam; PF00616; RasGAP; 1.
 DR SMART; SM00239; C2; 1.
 DR SMART; SM00323; RasGAP; 1.
 DR PROSITE; PS50018; RAS_GTPASE_ACTIV_2; 1.
 SQ SEQUENCE 967 AA; 106751 MW; 22536D86167137D4 CRC64;
 Query Match 46.5%; Score 153.5; DB 2; Length 967;
 Best Local Similarity 65.5%; Pred. No. 6.6e-10;
 Matches 36; Conservative 4; Mismatches 14; Indels 1; Gaps 1;
 QY 1 KGIISRLMSVEEELKRDHAEMQACGCGLTEDHLMPPRSFASFLDANVARLMSALT 55
 Db 907 KGIISRLMSVEEELKRDHAEMQACGCGLTEDHLMPPRSFASFLDANVARLMSALT 960
 RESULT 5
 Q6JTV1
 ID Q6JTV1 PRELIMINARY; PRT; 996 AA.
 AC Q6JTV1;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE DAB2IP.
 GN Name=Dab2ip;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=12956/SVEVtac; TISSUE=Spleen;
 RA Chen H., Karam J.A., Hsieh J.-T.;
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY305657; AAQ77380.1; --
 DR EMBL; AY305658; AAQ77381.1; --
 DR EMBL; AY305656; AAQ77379.1; --
 DR GO; GO:0005515; F:protein binding; IPI.
 DR InterPro; IPR000008; C2.
 DR InterPro; IPR008973; C2_CaLB.
 DR InterPro; IPR001936; RasGAP.
 DR Pfam; PF00168; C2; 1.
 DR Pfam; PF00616; RasGAP; 1.
 DR SMART; SM00239; C2; 1.
 DR SMART; SM00323; RasGAP; 1.
 DR PROSITE; PS50018; RAS_GTPASE_ACTIV_2; 1.
 SQ SEQUENCE 996 AA; 109994 MW; FD31F282048F9BF8 CRC64;
 Query Match 46.5%; Score 153.5; DB 2; Length 996;
 Best Local Similarity 65.5%; Pred. No. 6.8e-10;
 Matches 36; Conservative 4; Mismatches 14; Indels 1; Gaps 1;
 QY 1 KGIISRLMSVEEELKRDHAEMQACGCGLTEDHLMPPRSFASFLDANVARLMSALT 55
 Db 907 KGIISRLMSVEEELKRDHAEMQACGCGLTEDHLMPPRSFASFLDANVARLMSALT 960
 RESULT 6
 Q6P730
 ID Q6P730 PRELIMINARY; PRT; 996 AA.
 AC Q6P730;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Dab2ip protein.
 GN Name=Dab2ip;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Prostate;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek A.M., Gay L.J., Hulyk S.W.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Bouffard G.G.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Whiting J., Helton E., Kettman M., Madan A., Young A.C., Dickinson M.C.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Prostate;
 RA Strausberg R.;
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC061865; AAH61865.1; -;
 DR InterPro; IPR000008; C2.
 DR InterPro; IPR008973; C2_CaLB.
 DR InterPro; IPR001936; RasGAP.
 DR InterPro; IPR008936; Rho_GAP.
 DR Pfam; PF00168; C2; 1.
 DR Pfam; PF00616; RasGAP; 1.
 DR SMART; SM00239; C2; 1.
 DR SMART; SM00323; RasGAP; 1.
 DR PROSITE; PS50018; RAS_GTPASE_ACTIV_2; 1.
 DR PROSITE; PS50018; RAS_GTPASE_ACTIV_2; 1.
 SQ SEQUENCE 996 AA; 11004 MW; CA6B43D3129F4D5E CRC64;
 Query Match 46.5%; Score 153.5; DB 2; Length 996;
 Best Local Similarity 65.5%; Pred. No. 6.8e-10;
 Matches 36; Conservative 4; Mismatches 14; Indels 1; Gaps 1;
 QY 1 KGIISRLMSVEELKRDHAEMQAGCGLOTHLMPRRSAFASLDVAVNARLMSALT 55
 DB 907 KGIISRLMSVEELKRDHAEMQAGCGLOTHLMPRRSAFASLDVAVNARLMSALT 960
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Prostate;
 RA Strausberg R.;
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC061865; AAH61865.1; -;
 DR InterPro; IPR000008; C2.
 DR InterPro; IPR008973; C2_CaLB.
 DR InterPro; IPR001936; RasGAP.
 DR InterPro; IPR008936; Rho_GAP.
 DR Pfam; PF00168; C2; 1.
 DR Pfam; PF00616; RasGAP; 1.
 DR SMART; SM00239; C2; 1.
 DR SMART; SM00323; RasGAP; 1.
 DR PROSITE; PS50018; RAS_GTPASE_ACTIV_2; 1.
 DR PROSITE; PS50018; RAS_GTPASE_ACTIV_2; 1.
 SQ SEQUENCE 996 AA; 11004 MW; CA6B43D3129F4D5E CRC64;
 Query Match 46.5%; Score 153.5; DB 2; Length 996;
 Best Local Similarity 65.5%; Pred. No. 6.8e-10;
 Matches 36; Conservative 4; Mismatches 14; Indels 1; Gaps 1;
 QY 1 KGIISRLMSVEELKRDHAEMQAGCGLOTHLMPRRSAFASLDVAVNARLMSALT 55
 DB 907 KGIISRLMSVEELKRDHAEMQAGCGLOTHLMPRRSAFASLDVAVNARLMSALT 960
 RESULT 7
 Q924M9 ID Q924M9 PRELIMINARY; PRT; 996 AA.
 AC Q924M9;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE DOC2/DAB2 interactive protein.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Prostate;
 RX MEDLINE=21935348; PubMed=11812785;
 RA Wang Z., Teeng C.P., Pong R.C., Chen H., McConnell J.D., Navone N.,

RA Hsieh J.T.;
 RT "The mechanism of growth-inhibitory effect of DOC-2/DAB2 in prostate
 cancer. Characterization of a novel GTPase-activating protein
 associated with N-terminal domain of DOC-2/DAB2.";
 RL J. Biol. Chem. 277:12622-12631(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Prostate;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek A.M., Gay L.J., Hulyk S.W.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Bouffard G.G.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Whiting J., Helton E., Kettman M., Madan A., Young A.C., Dickinson M.C.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Prostate;
 RA Strausberg R.;
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF236130; AA93947.1; -;
 DR HSSP; P21359; INF1
 DR InterPro; IPR000008; C2.
 DR InterPro; IPR008973; C2_CaLB.
 DR InterPro; IPR001936; RasGAP.
 DR InterPro; IPR008936; Rho_GAP.
 DR Pfam; PF00168; C2; 1.
 DR Pfam; PF00616; RasGAP; 1.
 DR SMART; SM00239; C2; 1.
 DR SMART; SM00323; RasGAP; 1.
 DR PROSITE; PS50018; RAS_GTPASE_ACTIV_2; 1.
 DR PROSITE; PS50018; RAS_GTPASE_ACTIV_2; 1.
 SQ SEQUENCE 996 AA; 110075 MW; B4DF675548ED9EC CRC64;
 Query Match 46.5%; Score 153.5; DB 2; Length 996;
 Best Local Similarity 65.5%; Pred. No. 6.8e-10;
 Matches 36; Conservative 4; Mismatches 14; Indels 1; Gaps 1;
 QY 1 KGIISRLMSVEELKRDHAEMQAGCGLOTHLMPRRSAFASLDVAVNARLMSALT 55
 DB 907 KGIISRLMSVEELKRDHAEMQAGCGLOTHLMPRRSAFASLDVAVNARLMSALT 960
 RESULT 8
 AAH61865 ID AAH61865 PRELIMINARY; PRT; 996 AA.
 AC AAH61865;
 DT 14-APR-2004 (TrEMBLrel. 27, Created)
 DT 14-APR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 14-APR-2004 (TrEMBLrel. 27, Last annotation update)
 DE Dab2ip protein.
 GN DAB2IP.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Prostate;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek A.M., Gay L.J., Hulyk S.W.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Bouffard G.G.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Whiting J., Helton E., Kettman M., Madan A., Young A.C., Dickinson M.C.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Prostate;
 RA Strausberg R.;
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.

```

DR EMBL; BC061865; AAH61865.1; -.
SQ SEQUENCE 996 AA; 110004 MW; CAG6A3D3129F4D6E CRC64;

Query Match
Best Local Similarity 46.5%; Score 153.5; DB 2; Length 996;
Matches 36; Conservative 4; Mismatches 14; Indels 1; Gaps 1;

QY 1 KGIISRLMSVEEELKRDHAEQAGCGLOTEHLMPPRSAPASLDVAVNARLMSALT 55
|||||
DB 907 KGIISRLMSVEEELKRDHAEQAGCGLOTEHLMPPRSAPASLDVAVNARLMSALT 960

RESULT 9
AAQ77379
ID AAQ77379 PRELIMINARY; PRT; 996 AA.
AC AAQ77379;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DE DAB2IP.
GN DAB2IP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129S6/SvEvTac; TISSUE=Spleen;
RA Chen H., Karam J.A., Hsieh J.-T.;
RT "Cloning and characterization of mouse DAB2IP gene and its promoter.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY305656; AAQ77379.1; -.
SQ SEQUENCE 996 AA; 109994 MW; FD31F282048F9BF8 CRC64;

Query Match
Best Local Similarity 46.5%; Score 153.5; DB 2; Length 996;
Matches 36; Conservative 4; Mismatches 14; Indels 1; Gaps 1;

QY 1 KGIISRLMSVEEELKRDHAEQAGCGLOTEHLMPPRSAPASLDVAVNARLMSALT 55
|||||
DB 907 KGIISRLMSVEEELKRDHAEQAGCGLOTEHLMPPRSAPASLDVAVNARLMSALT 960

RESULT 10
AAQ77380
ID AAQ77380 PRELIMINARY; PRT; 996 AA.
AC AAQ77380;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DE DAB2IP.
GN DAB2IP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129S6/SvEvTac; TISSUE=Spleen;
RA Chen H., Karam J.A., Hsieh J.-T.;
RT "Cloning and characterization of mouse DAB2IP gene and its promoter.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY305657; AAQ77380.1; -.
SQ SEQUENCE 996 AA; 109994 MW; FD31F282048F9BF8 CRC64;

Query Match
Best Local Similarity 46.5%; Score 153.5; DB 2; Length 996;
Matches 36; Conservative 4; Mismatches 14; Indels 1; Gaps 1;

QY 1 KGIISRLMSVEEELKRDHAEQAGCGLOTEHLMPPRSAPASLDVAVNARLMSALT 55
|||||
DB 907 KGIISRLMSVEEELKRDHAEQAGCGLOTEHLMPPRSAPASLDVAVNARLMSALT 960

```

```

RESULT 11
AAQ77381
ID AAQ77381 PRELIMINARY; PRT; 996 AA.
AC AAQ77381;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DE DAB2IP.
GN DAB2IP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129S6/SvEvTac; TISSUE=Spleen;
RA Chen H., Karam J.A., Hsieh J.-T.;
RT "Cloning and characterization of mouse DAB2IP gene and its promoter.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY305658; AAQ77381.1; -.
SQ SEQUENCE 996 AA; 109994 MW; FD31F282048F9BF8 CRC64;

Query Match
Best Local Similarity 46.5%; Score 153.5; DB 2; Length 996;
Matches 36; Conservative 4; Mismatches 14; Indels 1; Gaps 1;

QY 1 KGIISRLMSVEEELKRDHAEQAGCGLOTEHLMPPRSAPASLDVAVNARLMSALT 55
|||||
DB 907 KGIISRLMSVEEELKRDHAEQAGCGLOTEHLMPPRSAPASLDVAVNARLMSALT 960

RESULT 12
Q9C0C0
ID Q9C0C0 PRELIMINARY; PRT; 1036 AA.
AC Q9C0C0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE KIAA1743 protein (fragment).
GN Name=KIAA1743;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=21082932; PubMed=11214970;
RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIX.
The complete sequences of 100 new cDNA clones from brain which code
for large proteins in vitro.";
RL DNA Res. 7:347-355(2000).
DR EMBL; AB051530; BAB21834.2; -.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2_CaLB.
DR InterPro; IPR001849; PH.
DR InterPro; IPR011036; PH_related.
DR InterPro; IPR01936; RasGAP.
DR InterPro; IPR008936; Rho_GAP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00616; RasGAP; 1.
DR SMART; SM00239; C2; 1.
DR SMART; SM00323; RasGAP; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS50018; RAS_GTPASE_ACTIV_2; 1.
FT NON_TER
SQ SEQUENCE 1036 AA; 114409 MW; 1B07D8D3B9DEA76A CRC64;

Query Match
Best Local Similarity 46.5%; Score 153.5; DB 2; Length 1036;
Matches 65.5%; Pred. No. 7.1e-10;

```

Matches 36; Conservative 4; Mismatches 14; Indels 1; Gaps 1;

QY 1 KGIISRLMSVEELKRDHAEKQAGCGLTQEDHLMPPRSFAFASLDVAVNARLMSALT 55
 DB 976 KGIISRLMSVEELKRDHAEKQAGCGLTQEDHLMPPRSFAFASLDVAVNARLMSALT 1029

RESULT 13

Q96SE1 PRELIMINARY; PRT; 1065 AA.

AC Q96SE1 (TREMELrel. 19, Created)
 DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE NGAP-like protein.
 GN Name=AF9q34;
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA von Bergh A.R.M., Wijers-Koster P.M., Groot A.J., Kluin P.M.,
 RA Schuring E.
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY032952; AAK50336.1; -.
 DR HSSP; P21359; INF1.
 DR Genew; HGNC:17294; DAB2IP.
 DR InterPro; IPR000008; C2.
 DR InterPro; IPR008973; C2_CaLB.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR011036; PH_related.
 DR InterPro; IPR001936; RasGAP.
 DR InterPro; IPR008936; Rho_GAP.
 DR Pfam; PF00168; C2; 1.
 DR Pfam; PF00169; PH; 1.
 DR Pfam; PF00616; RasGAP; 1.
 DR SMART; SM00239; C2; 1.
 DR SMART; SM00323; RasGAP; 1.
 DR SMART; SM00323; RasGAP; 1.
 DR PROSITE; PS00499; C2_DOMAIN_1; FALSE_NEG.
 DR PROSITE; PS00004; C2_DOMAIN_2; FALSE_NEG.
 DR PROSITE; PS50003; PH_DOMAIN; FALSE_NEG.
 DR PROSITE; PS50009; RAS_GTPASE_ACTIV_1; FALSE_NEG.
 DR PROSITE; PS50018; RAS_GTPASE_ACTIV_2; 1.
 KW GTPase activation.
 FT DOMAIN 41 158 PH.
 FT DOMAIN 170 266 C2 domain.
 FT DOMAIN 327 519 Ras-GAP.
 FT DOMAIN 237 243 Poly-Lys.
 FT DOMAIN 922 925 Poly-Ser.
 SQ SEQUENCE 1139 AA; 128556 MW; 4787F41F02108AA3 CRC64;

Query Match 46.5%; Score 153.5; DB 2; Length 1065;
 Best Local Similarity 65.5%; Pred. No. 7.3e-10;
 Matches 36; Conservative 4; Mismatches 14; Indels 1; Gaps 1;

QY 1 KGIISRLMSVEELKRDHAEKQAGCGLTQEDHLMPPRSFAFASLDVAVNARLMSALT 55
 DB 976 KGIISRLMSVEELKRDHAEKQAGCGLTQEDHLMPPRSFAFASLDVAVNARLMSALT 1029

RESULT 14

NGAP_HUMAN STANDARD; PRT; 1139 AA.

AC NGAP_HUMAN
 ID Q9UJF2; Q95174;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Ras GTPase-activating protein NGAP (RAS protein activator like 1).
 GN Name=RASAL2; Synonyms=NGAP;
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE=Heart;
 RC MEDLINE=99093006; PubMed=9877179;
 RA Noto S., Maeda T., Hattori S., Inazawa J., Imamura M., Asaka M.,
 RA Hatakeyama M.
 RT "A novel human RasGAP-like gene that maps within the prostate cancer
 susceptibility locus at chromosome 1q25.";

FEBS Lett. 441:127-131 (1998).
 [2]
 RN SEQUENCE OF 41-1139 FROM N.A.
 RP Chapman J.;
 RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Inhibitory regulator of the Ras-cyclic AMP pathway.
 CC -!- SIMILARITY: Contains 1 C2 domain.
 CC -!- SIMILARITY: Contains 1 PH domain.
 CC -!- SIMILARITY: Contains 1 Ras-GAP domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF047711; AAD04814.1; -.
 CC EMBL; AL035702; CAB53260.1; -.
 CC HSSP; P21359; INF1.
 CC IntAct; Q9UJF2; -.
 CC Genew; HGNC:9874; RASAL2.
 CC MIM; 606136; -.
 CC GO; GO:0005099; P:Ras GTPase activator activity; TAS.
 CC GO; GO:0007165; P:signal transduction; TAS.
 CC InterPro; IPR000008; C2.
 CC InterPro; IPR008973; C2_CaLB.
 CC InterPro; IPR001849; PH.
 CC InterPro; IPR011036; PH_related.
 CC InterPro; IPR001936; RasGAP.
 CC InterPro; IPR008936; Rho_GAP.
 CC Pfam; PF00168; C2; 1.
 CC Pfam; PF00616; RasGAP; 1.
 CC SMART; SM00239; C2; 1.
 CC SMART; SM00323; PH; 1.
 CC SMART; SM00323; RasGAP; 1.
 CC PROSITE; PS00499; C2_DOMAIN_1; FALSE_NEG.
 CC PROSITE; PS00004; C2_DOMAIN_2; FALSE_NEG.
 CC PROSITE; PS50003; PH_DOMAIN; FALSE_NEG.
 CC PROSITE; PS50009; RAS_GTPASE_ACTIV_1; FALSE_NEG.
 CC PROSITE; PS50018; RAS_GTPASE_ACTIV_2; 1.
 KW GTPase activation.
 FT DOMAIN 41 158 PH.
 FT DOMAIN 170 266 C2 domain.
 FT DOMAIN 327 519 Ras-GAP.
 FT DOMAIN 237 243 Poly-Lys.
 FT DOMAIN 922 925 Poly-Ser.
 SQ SEQUENCE 1139 AA; 128556 MW; 4A65C8243E1259A3 CRC64;

Query Match 39.8%; Score 131.5; DB 1; Length 1139;
 Best Local Similarity 56.4%; Pred. No. 4.1e-07;
 Matches 31; Conservative 7; Mismatches 16; Indels 1; Gaps 1;

QY 1 KGIISRLMSVEELKRDHAEKQAGCGLTQEDHLMPPRSFAFASLDVAVNARLMSALT 55
 DB 1051 KSIISRLMAVEELKRDHAEKQAGCGLTQEDHLMPPRSFAFASLDVAVNARLMSALT 1104

RESULT 15

GAP2_DROME STANDARD; PRT; 1556 AA.

AC GAP2_DROME
 ID Q8T498; Q95TL0; Q9VX23;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Probable Ras GTPase-activating protein.
 GN ORFNames=CG32560;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RA SEQUENCE FROM N.A.
 RA STRAIN=Berkeley;
 RA MEDLINE=20196066; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scher S.E., Li P.W., Hoskins A.A., Calle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers J.H., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.C., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Aril J.F., Agayuni A., An H.-J., Andrews-Pfannkoch C., Balgwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Fostel C., Gabriellian A.E., Garr N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Hostin N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin K.A., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Moberg C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Renington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E.C., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RA "the genome sequence of Drosophila melanogaster";
 RA Science 287:2185-2195(2000).
 [2]
 RA REVISIONS.
 RA MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kinkner J.S., Millburn G.H., Prochuk S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RA "Annotation of the Drosophila melanogaster euchromatic genome: a
 RA systematic review";
 RA Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 [3]
 RA SEQUENCE FROM N.A.
 RA STRAIN=Berkeley; TISSUE=Embryo, and Testis;
 RA MEDLINE=22426066; PubMed=12537569;
 RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
 RA George R.A., Guarin H., Krommiller B., Pacleb J.M., Park S., Wan K.H.,
 RA Rubin G.M., Celniker S.E.;
 RA "A Drosophila full-length cDNA resource";
 RA Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
 CC -!- FUNCTION: Probable GTPase-activating protein, which may act as a
 CC negative regulator for some member of the Ras family. They
 CC decrease the signaling activity of Ras by stimulating its intrinsic
 CC GTPase activity, thereby lowering the levels of GTP-bound, active
 CC Ras (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Contains 1 C2 domain.
 CC -!- SIMILARITY: Contains 1 PH domain.
 CC -!- SIMILARITY: Contains 1 Ras-GAP domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AE003506; AAF48759.2; -
 CC EMBL; AY058706; AAL13935.1; -
 CC EMBL; AY089290; AAL90028.1; ALT_INIT.
 CC InAct; Q8T498; -
 CC FLYbase; FBgn0052560; CG32560.
 CC GO; GO:0005737; C:cytoplasm; ISS.
 CC GO; GO:0005039; F:Ras GTPase activator activity; ISS.
 CC GO; GO:0046580; P:negative regulation of Ras protein signal t.; ISS.
 CC InterPro; IPR000008; C2.
 CC InterPro; IPR008973; C2_CaLB.
 CC InterPro; IPR001849; PH_related.
 CC InterPro; IPR011036; PH_related.
 CC InterPro; IPR001936; RasGAP.
 CC InterPro; IPR008936; Rho_GAP.
 CC Pfam; PF00169; C2; 1.
 CC Pfam; PF00169; PH; 1.
 CC Pfam; PF00616; RasGAP; 1.
 CC SMART; SM00239; C2; 1.
 CC SMART; SM00233; PH; 1.
 CC SMART; SM00323; RasGAP; 1.
 CC PROSITE; PS00499; C2 DOMAIN 1; FALSE NEG.
 CC PROSITE; PS00004; C2 DOMAIN 2; FALSE NEG.
 CC PROSITE; PS00003; PH DOMAIN; FALSE NEG.
 CC PROSITE; PS00509; RAS GTPASE ACTIV_1; FALSE NEG.
 CC PROSITE; PS00018; RAS_GTPASE_ACTIV_2; 1.
 CC GTPase activation.
 CC DOMAIN 267 PH.
 CC FT DOMAIN 209 267
 CC FT DOMAIN 279 360 C2 domain.
 CC FT DOMAIN 450 644 Ras-GAP.
 CC FT DOMAIN 346 349 Poly-Lys.
 CC FT DOMAIN 392 397 Poly-Gly.
 CC FT DOMAIN 835 856 Poly-Gln.
 CC FT DOMAIN 1183 1192 Poly-Gln.
 CC FT DOMAIN 1241 1248 Poly-Ala.
 CC FT DOMAIN 1264 1275 Poly-Ser.
 CC FT DOMAIN 1328 1331 Poly-Gln.
 CC FT DOMAIN 1335 1339 Poly-Gln.
 CC FT DOMAIN 1403 1411 Poly-Gly.
 CC FT DOMAIN 1540 1547 Poly-Gln.
 CC FT CONFLICT 1346 1346 H -> P (in Ref. 3; AAL90028).
 CC SEQUENCE 1556 AA; 170797 MW; 1EA002B6DFF7D7F8 CRC64;
 Query Match 27.4%; Score 90.5; DB 1; Length 1556;
 Best Local Similarity 39.3%; Pred. No. 0.067;
 Matches 22; Conservative 14; Mismatches 17; Indels 3; Gaps 2;
 QY 1 KGIISRLMSVEELKRDHAEMQACGLQTEDHLMPPR-SAFASLDVAVNRLMSALT 55
 DB 1478 RSIIDRLITMEELRREQLKM--SLALSHKQVIEBQGGQIQALDAANSRLSALT 1531
 Search completed: November 10, 2004, 12:27:24
 Job time : 75.3071 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:41:17 ; Search time 7.74383 Seconds
(without alignments)
222.664 Million cell updates/sec

Title: US-10-092-750-21

Perfect score: 139
Sequence: 1 WERIEERLAYIADHLGFSWTELARAL 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

```
Database :
1: /cgn2_6/pdata/1/iaa/5A_COMB pep.*
2: /cgn2_6/pdata/1/iaa/5B_COMB pep.*
3: /cgn2_6/pdata/1/iaa/6A_COMB pep.*
4: /cgn2_6/pdata/1/iaa/6B_COMB pep.*
5: /cgn2_6/pdata/1/iaa/PCTUS_COMB.pe
6: /cgn2_6/pdata/1/iaa/backfiles1.pe
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query | | DB | ID | Description |
|------------|-------|-------|--------|----|----------------------|--------------------|
| | | Match | Length | | | |
| 1 | 123 | 88.5 | 1839 | 2 | US-09-172-977-4 | Sequence 4, Appli |
| 2 | 123 | 88.5 | 1839 | 4 | US-09-404-108-4 | Sequence 4, Appli |
| 3 | 123 | 88.5 | 3924 | 4 | US-09-538-098-1246 | Sequence 1246, Ap |
| 4 | 52 | 37.4 | 835 | 4 | US-09-619-353-7 | Sequence 7, Appli |
| 5 | 49.5 | 35.6 | 51 | 1 | US-08-519-103-16 | Sequence 16, Appli |
| 6 | 49.5 | 35.6 | 51 | 1 | US-08-519-103-18 | Sequence 18, Appli |
| 7 | 49.5 | 35.6 | 51 | 3 | US-09-018-635-16 | Sequence 16, Appli |
| 8 | 49.5 | 35.6 | 51 | 3 | US-09-018-635-18 | Sequence 18, Appli |
| 9 | 49.5 | 35.6 | 51 | 4 | US-09-912-962-16 | Sequence 16, Appli |
| 10 | 49.5 | 35.6 | 51 | 4 | US-09-912-962-18 | Sequence 18, Appli |
| 11 | 49.5 | 35.6 | 156 | 3 | US-08-928-941D-4 | Sequence 4, Appli |
| 12 | 49.5 | 35.6 | 156 | 3 | US-08-928-941D-36 | Sequence 36, Appli |
| 13 | 49.5 | 35.6 | 156 | 3 | US-09-280-590A-4 | Sequence 4, Appli |
| 14 | 49.5 | 35.6 | 156 | 3 | US-09-280-590A-46 | Sequence 46, Appli |
| 15 | 49.5 | 35.6 | 156 | 4 | US-09-892-398-4 | Sequence 4, Appli |
| 16 | 49.5 | 35.6 | 156 | 4 | US-09-892-398-46 | Sequence 46, Appli |
| 17 | 49.5 | 35.6 | 640 | 4 | US-09-538-093-891 | Sequence 891, App |
| 18 | 49.5 | 35.6 | 751 | 4 | US-09-402-929-2 | Sequence 2, Appli |
| 19 | 49.5 | 35.6 | 752 | 4 | US-09-402-929-6 | Sequence 6, Appli |
| 20 | 49.5 | 35.6 | 752 | 4 | US-09-538-092-892 | Sequence 892, App |
| 21 | 48.5 | 34.9 | 414 | 5 | PCT-US93-03077-3 | Sequence 3, Appli |
| 22 | 48.5 | 34.9 | 1093 | 5 | PCT-US93-03077-1 | Sequence 1, Appli |
| 23 | 48 | 34.5 | 138 | 4 | US-09-270-767-47463 | Sequence 47463, A |
| 24 | 48 | 34.5 | 424 | 4 | US-09-248-796A-16232 | Sequence 16232, A |
| 25 | 47 | 33.8 | 68 | 6 | 5320958-11 | Patent No. 5320958 |
| 26 | 47 | 33.8 | 266 | 2 | US-08-845-161A-6 | Sequence 6, Appli |
| 27 | 47 | 33.8 | 266 | 3 | US-09-270-751-6 | Sequence 6, Appli |

```
; LENGTH: 1839
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: g29491
US-09-404-108-4

Query Match      88.5%; Score 123; DB 4; Length 1839;
Best Local Similarity 96.0%; Pred. No. 6.4e-10;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ERIEERLAYIADHGLGFSWTELARAL 26
Db 1449 ERIEERLAYIADHGLGFSWTELARAL 1473

RESULT 3
US-09-538-092-1246
; Sequence 1246, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSeqFormatter Version 0.9
; SEQ ID NO 1246
; LENGTH: 3924
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number Q01484
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number Q01485
US-09-538-092-1246

Query Match      88.5%; Score 123; DB 4; Length 3924;
Best Local Similarity 96.0%; Pred. No. 1.5e-09;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ERIEERLAYIADHGLGFSWTELARAL 26
Db 3534 ERIEERLAYIADHGLGFSWTELARAL 3558

RESULT 4
US-09-619-353-7
; Sequence 7, Application US/09619353
; Patent No. 6410249
; GENERAL INFORMATION:
; APPLICANT: Ngai, John
; APPLICANT: Speca, David J.
; APPLICANT: Lin, David M.
; APPLICANT: Isaacoff, Ehud Y.
; APPLICANT: Dittman, Andrew H.
; APPLICANT: Fan, Jinhong
; TITLE OF INVENTION: Odorant Receptors
; FILE REFERENCE: B99-038-2
; CURRENT APPLICATION NUMBER: US/09/619,353
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 60/144,766
; PRIOR FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 14

; LENGTH: 835
; TYPE: PRT
; ORGANISM: Carassius auratus
; OTHER INFORMATION: g29491
US-09-619-353-7

Query Match      37.4%; Score 52; DB 4; Length 835;
Best Local Similarity 75.0%; Pred. No. 15;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 LAYIADHGLGFSW 19
Db 199 LAYIVKHLGWSW 210

RESULT 5
US-08-519-103-16
; Sequence 16, Application US/08519103
; Patent No. 5733730
; GENERAL INFORMATION:
; APPLICANT: deLange, Titia
; TITLE OF INVENTION: TELOMERE REPEAT BINDING FACTOR AND
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KLAUBER & JACKSON
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/519,103
; APPLICATION NUMBER: US/08/519,103
; FILING DATE: 25-AUG-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Reury, Sharon E.
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 600-1-142
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 51 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-519-103-16

Query Match      35.6%; Score 49.5; DB 1; Length 51;
Best Local Similarity 40.7%; Pred. No. 1.6;
Matches 11; Conservative 4; Mismatches 11; Indels 1; Gaps 1;

QY 1 WERIEERLAYIA-DHGLGFSWTELARAL 26
Db 6 WTEEDRIIYEAKKRLGNRWAEIAKLL 32

RESULT 6
US-08-519-103-18
; Sequence 18, Application US/08519103
; Patent No. 5733730
; GENERAL INFORMATION:
; APPLICANT: deLange, Titia
```



```
; TITLE OF INVENTION: TELOMERE REPEAT BINDING FACTOR AND
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC USE THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KLAUBER & JACKSON
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/519,103
; FILING DATE: 25-AUG-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon E.
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 600-1-142
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 51 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-519-103-18
;
; Query Match 35.6%; Score 49.5; DB 1; Length 51;
; Best Local Similarity 40.7%; Pred. No. 1.6;
; Matches 11; Conservative 4; Mismatches 11; Indels 1; Gaps 1;
;
; Qy 1 WERIEERLAYIA-DHLGFSWTELARAL 26
; Db 6 WTEEDRIIYQAHKRLGNRWAEIAKLL 32
;
; RESULT 7
; US-09-018-635-16
; Sequence 16, Application US/09018635
; Patent No. 6297356
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Broccoli, Dominique
; APPLICANT: Smogorzewska, Agata
; TITLE OF INVENTION: TELOMERE REPEAT BINDING FACTOR AND
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC USE THEREOF
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KLAUBER & JACKSON
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/018,635
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: David A. Jackson
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-142 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 51 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-018-635-16
;
; Query Match 35.6%; Score 49.5; DB 3; Length 51;
; Best Local Similarity 40.7%; Pred. No. 1.6;
; Matches 11; Conservative 4; Mismatches 11; Indels 1; Gaps 1;
```

```
; NAME: David A. Jackson
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-142 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 51 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-018-635-16
;
; Query Match 35.6%; Score 49.5; DB 3; Length 51;
; Best Local Similarity 40.7%; Pred. No. 1.6;
; Matches 11; Conservative 4; Mismatches 11; Indels 1; Gaps 1;
;
; Qy 1 WERIEERLAYIA-DHLGFSWTELARAL 26
; Db 6 WTEEDRIIYQAHKRLGNRWAEIAKLL 32
;
; RESULT 8
; US-09-018-635-18
; Sequence 18, Application US/09018635
; Patent No. 6297356
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Broccoli, Dominique
; APPLICANT: Smogorzewska, Agata
; TITLE OF INVENTION: TELOMERE REPEAT BINDING FACTOR AND
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC USE THEREOF
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KLAUBER & JACKSON
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/018,635
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: David A. Jackson
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-142 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 51 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-018-635-18
;
; Query Match 35.6%; Score 49.5; DB 3; Length 51;
; Best Local Similarity 40.7%; Pred. No. 1.6;
; Matches 11; Conservative 4; Mismatches 11; Indels 1; Gaps 1;
```



```

; ORGANISM: Gallus gallus
; US-08-928-941D-36
;
; Query Match 35.6%; Score 49.5; DB 3; Length 156;
; Best Local Similarity 40.7%; Pred. No. 5.6;
; Matches 11; Conservative 4; Mismatches 11; Indels 1; Gaps 1;
;
Qy 1 WERIEERLAYIA-DHLGFSWTELARAL 26
Db 110 WTEEDRIIYQAHKRLGNRWAETAKLL 136
;
RESULT 13
US-09-280-590A-4
; Sequence 4, Application US/09280590A
; Patent No. 6303772
; GENERAL INFORMATION:
; APPLICANT: Hirai, Hiroshi
; Sherr, Charles
; Inoue, Kazushi
; Bodner, Sarah M.
;
; TITLE OF INVENTION: CYCLIN-D BINDING FACTOR, AND USES
; THEREOF
;
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/280,590A
; FILING DATE: 29-Mar-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1340-1-002 N CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
;
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 156 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
; US-09-280-590A-4
;
; Query Match 35.6%; Score 49.5; DB 3; Length 156;
; Best Local Similarity 40.7%; Pred. No. 5.6;
; Matches 11; Conservative 4; Mismatches 11; Indels 1; Gaps 1;
;
Qy 1 WERIEERLAYIA-DHLGFSWTELARAL 26
Db 110 WTEEDRIIYQAHKRLGNRWAETAKLL 136
;
RESULT 14
US-09-280-590A-46

```

; Sequence 46, Application US/09280590A
; Patent No. 6303772
; GENERAL INFORMATION:
; APPLICANT: Hirai, Hiroshi
; Sherr, Charles
; Inoue, Kazushi
; Bodner, Sarah M.
; TITLE OF INVENTION: CYCLIN-D BINDING FACTOR, AND USES
; THEREOF
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/280,590A
; FILING DATE: 29-Mar-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1340-1-002 N CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 156 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; FRAGMENT TYPE: <Unknown>
; ORIGINAL SOURCE:
; ORGANISM: Gallus gallus
; SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-09-280-590A-46
Query Match 35.6%; Score 49.5; DB 3; Length 156;
Best Local Similarity 40.7%; Pred. No. 5.6;
Matches 11; Conservative 4; Mismatches 11; Indels 1; Gaps 1;
Qy 1 WERIEERLAYIA-DHLGFSWTELARAL 26
Db 110 WTEEDRIIYQAHKRLGNRWAEIAKLL 136
RESULT 15
US-09-892-398-4
; Sequence 4, Application US/09892398
; Patent No. 6673902
; GENERAL INFORMATION:
; APPLICANT: Hirai, Hiroshi
; Sherr, Charles
; Inoue, Kazushi
; Bodner, Sarah M.
; TITLE OF INVENTION: CYCLIN-D BINDING FACTOR, AND USES
; THEREOF
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; Floor

; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/892,398
; FILING DATE: 27-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/280,590
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1340-1-002 N CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 156 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-892-398-4
Query Match 35.6%; Score 49.5; DB 4; Length 156;
Best Local Similarity 40.7%; Pred. No. 5.6;
Matches 11; Conservative 4; Mismatches 11; Indels 1; Gaps 1;
Qy 1 WERIEERLAYIA-DHLGFSWTELARAL 26
Db 110 WTEEDRIIYQAHKRLGNRWAEIAKLL 136
Search completed: November 10, 2004, 12:32:32
Job time : 7.74383 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 15:53:52 ; Search time 25.0772 Seconds
(without alignments)
366.225 Million cell updates/sec

Title: US-10-092-750-21

Perfect score: 139

Sequence: 1 WERIEERLAYIADHLGFSWTELARAL 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/prodata1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/prodata1/pubpaa/US06_PUBCOMB.pep.*
- 4: /cgn2_6/prodata1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/prodata1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/prodata1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/prodata1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/prodata1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/prodata1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/prodata1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/prodata1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/prodata1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/prodata1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/prodata1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/prodata1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/prodata1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/prodata1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/prodata1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/prodata1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/prodata1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------------------|-------------------|
| 1 | 139 | 100.0 | 26 | 14 US-10-092-750-21 | Sequence 21, Appl |
| 2 | 123 | 88.5 | 396 | 14 US-10-336-031-1 | Sequence 1, Appl |
| 3 | 112 | 80.6 | 253 | 17 US-10-425-115-229829 | Sequence 229829, |
| 4 | 90 | 64.7 | 294 | 14 US-10-336-031-2 | Sequence 2, Appl |
| 5 | 90 | 64.7 | 400 | 14 US-10-104-047-2995 | Sequence 2995, Ap |
| 6 | 90 | 64.7 | 1762 | 14 US-10-203-194-117 | Sequence 117, App |
| 7 | 90 | 64.7 | 3913 | 15 US-10-334-143-45 | Sequence 45, Appl |
| 8 | 57.5 | 41.4 | 95 | 11 US-09-972-211-128 | Sequence 128, App |
| 9 | 57.5 | 41.4 | 95 | 15 US-10-087-684-87 | Sequence 87, Appl |
| 10 | 57.5 | 41.4 | 95 | 15 US-10-218-779-87 | Sequence 87, Appl |
| 11 | 57.5 | 41.4 | 95 | 15 US-10-096-625-128 | Sequence 128, App |
| 12 | 57.5 | 41.4 | 96 | 10 US-09-970-944-45 | Sequence 45, Appl |
| 13 | 57.5 | 41.4 | 96 | 15 US-10-037-417-124 | Sequence 124, App |

| | | | | | | |
|----|------|------|------|----|----------------------|-------------------|
| 14 | 56 | 40.3 | 384 | 14 | US-10-156-761-9326 | Sequence 9326, Ap |
| 15 | 52 | 37.4 | 835 | 14 | US-10-151-208-7 | Sequence 7, Appl |
| 16 | 51.5 | 37.1 | 278 | 17 | US-10-729-930-9645 | Sequence 9645, Ap |
| 17 | 51.5 | 37.1 | 488 | 14 | US-10-156-761-7619 | Sequence 7619, Ap |
| 18 | 51 | 36.7 | 857 | 14 | US-10-156-761-12929 | Sequence 12929, A |
| 19 | 50 | 36.0 | 557 | 14 | US-10-369-493-4065 | Sequence 4065, Ap |
| 20 | 49.5 | 35.6 | 51 | 9 | US-09-912-962-16 | Sequence 16, Appl |
| 21 | 49.5 | 35.6 | 51 | 9 | US-09-912-962-18 | Sequence 18, Appl |
| 22 | 49.5 | 35.6 | 156 | 10 | US-09-892-398-4 | Sequence 4, Appl |
| 23 | 49.5 | 35.6 | 156 | 10 | US-09-892-398-46 | Sequence 46, Appl |
| 24 | 49 | 35.3 | 44 | 14 | US-10-321-857-24 | Sequence 24, Appl |
| 25 | 49 | 35.3 | 44 | 14 | US-10-318-675-24 | Sequence 24, Appl |
| 26 | 48.5 | 34.9 | 224 | 14 | US-10-156-761-11392 | Sequence 11392, A |
| 27 | 48.5 | 34.9 | 1069 | 14 | US-10-146-473-77 | Sequence 77, Appl |
| 28 | 48.5 | 34.9 | 1093 | 16 | US-10-408-765A-1961 | Sequence 1961, Ap |
| 29 | 48 | 34.5 | 310 | 15 | US-10-424-599-163497 | Sequence 163497, |
| 30 | 48 | 34.5 | 988 | 14 | US-10-369-493-6548 | Sequence 6548, Ap |
| 31 | 47.5 | 34.2 | 234 | 14 | US-10-156-761-9923 | Sequence 9923, Ap |
| 32 | 47.5 | 34.2 | 300 | 16 | US-10-437-963-167029 | Sequence 167029, |
| 33 | 47.5 | 34.2 | 717 | 15 | US-10-282-122A-77817 | Sequence 77817, A |
| 34 | 47 | 33.8 | 68 | 8 | US-08-808-031A-11 | Sequence 11, Appl |
| 35 | 47 | 33.8 | 142 | 15 | US-10-424-599-267724 | Sequence 267724, |
| 36 | 47 | 33.8 | 241 | 8 | US-08-808-031A-33 | Sequence 33, Appl |
| 37 | 47 | 33.8 | 243 | 8 | US-08-808-031A-32 | Sequence 32, Appl |
| 38 | 47 | 33.8 | 266 | 8 | US-08-808-031A-6 | Sequence 6, Appl |
| 39 | 47 | 33.8 | 266 | 14 | US-10-373-877-6 | Sequence 6, Appl |
| 40 | 47 | 33.8 | 480 | 8 | US-08-808-031A-42 | Sequence 42, Appl |
| 41 | 47 | 33.8 | 485 | 8 | US-08-808-031A-2 | Sequence 2, Appl |
| 42 | 46.5 | 33.5 | 38 | 14 | US-10-351-641-1078 | Sequence 1078, Ap |
| 43 | 46.5 | 33.5 | 130 | 16 | US-10-767-701-38100 | Sequence 38100, A |
| 44 | 46.5 | 33.5 | 131 | 14 | US-10-021-811-60 | Sequence 60, Appl |
| 45 | 46.5 | 33.5 | 131 | 15 | US-10-859-869-60 | Sequence 60, Appl |

ALIGNMENTS

RESULT 1
US-10-092-750-21
; Sequence 21, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-XL
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092.750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-21

Query Match 100.0%; Score 139; DB 14; Length 26;
Best Local Similarity 100.0%; Pred. No. 3.6e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WERIEERLAYIADHLGFSWTELARAL 26
DB 1 WERIEERLAYIADHLGFSWTELARAL 26

RESULT 2
US-10-336-031-1
; Sequence 1, Application US/10336031
; Publication No. US20030167491A1
; GENERAL INFORMATION:

```
; APPLICANT: BENNETT, VANN
; APPLICANT: GRAMOLINI, ANTHONY O.
; APPLICANT: MOHLER, PETER J.
; TITLE OF INVENTION: METHODS OF MODULATING LOCALIZATION AND PHYSIOLOGICAL
; TITLE OF INVENTION: FUNCTION OF IP3 RECEPTORS
; FILE REFERENCE: 1579-777
; CURRENT APPLICATION NUMBER: US/10/336,031
; CURRENT FILING DATE: 2003-01-03
; PRIOR FILING DATE: 2002-01-03
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-336-031-1

Query Match      88.5%; Score 123; DB 14; Length 396;
Best Local Similarity 96.0%; Pred. No. 1.5e-09;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ERIBERLAYIADHLGFSWTELARAL 26
|||:|||||:|||||:|||||:
Db 6 ERIBERLAYIADHLGFSWTELAREL 30
|||:|||||:|||||:|||||:

RESULT 3
US-10-425-115-229829
; Sequence 229829, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 229829
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(253)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_141200C.1.pep
US-10-425-115-229829

Query Match      80.6%; Score 112; DB 17; Length 253;
Best Local Similarity 95.5%; Pred. No. 3.8e-08;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ERIBERLAYIADHLGFSWTELA 23
|||:|||||:|||||:|||||:
Db 232 ERMEERLAYIADHLGFSWTELA 253
|||:|||||:|||||:|||||:

RESULT 4
US-10-336-031-2
; Sequence 2, Application US/10336031
; Publication No. US20030167491A1
; GENERAL INFORMATION:
; APPLICANT: BENNETT, VANN
; APPLICANT: GRAMOLINI, ANTHONY O.
; APPLICANT: MOHLER, PETER J.
; TITLE OF INVENTION: METHODS OF MODULATING LOCALIZATION AND PHYSIOLOGICAL
; TITLE OF INVENTION: FUNCTION OF IP3 RECEPTORS
```

```
; FILE REFERENCE: 1579-777
; CURRENT APPLICATION NUMBER: US/10/336,031
; CURRENT FILING DATE: 2003-01-03
; PRIOR FILING DATE: 2002-01-03
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-336-031-2

Query Match      64.7%; Score 90; DB 14; Length 294;
Best Local Similarity 68.0%; Pred. No. 7.4e-05;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 ERIBERLAYIADHLGFSWTELARAL 26
|||:|||||:|||||:|||||:
Db 6 ERTDIRMAIVADHLGSLWTELAREL 30
|||:|||||:|||||:|||||:

RESULT 5
US-10-104-047-2995
; Sequence 2995, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1el full length cdna
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2995
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2995

Query Match      64.7%; Score 90; DB 14; Length 400;
Best Local Similarity 68.0%; Pred. No. 0.0001;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 ERIBERLAYIADHLGFSWTELARAL 26
|||:|||||:|||||:|||||:
Db 8 ERTDIRMAIVADHLGSLWTELAREL 32
|||:|||||:|||||:|||||:

RESULT 6
US-10-205-194-117
; Sequence 117, Application US/10205194
; Publication No. US20030134301A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pinnock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018201
; CURRENT APPLICATION NUMBER: US/10/205,194
; CURRENT FILING DATE: 2000-07-24
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 117
; LENGTH: 1762
; TYPE: PRT
; ORGANISM: Rattus norvegicus
```

```

; FEATURE:
; OTHER INFORMATION: Ankyrin isoform
US-10-205-194-117

Query Match      64.7%  Score 90; DB 14; Length 1762;
Best Local Similarity 68.0%; Pred. No. 0.00052;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 ERBERLAYIADHLGFSWTELARAL 26
Db 1474 ERTDIRMAIVADHLGLSWTELAREL 1498
      ||:|:|||||
RESULT 7
US-10-334-143-45
; Sequence 45, Application US/10334143
; Publication No. US20040009549A1
; GENERAL INFORMATION:
; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH
; APPLICANT: SUDASANAM, SUCHA
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL
; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD
; FILE REFERENCE: 038602/1543
; CURRENT APPLICATION NUMBER: US/10/334,143
; PRIOR FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: 60/343,169
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 3913
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-334-143-45

Query Match      64.7%  Score 90; DB 15; Length 3913;
Best Local Similarity 68.0%; Pred. No. 0.0012;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 ERBERLAYIADHLGFSWTELARAL 26
Db 3624 ERTDIRMAIVADHLGLSWTELAREL 3648
      ||:|:|||||
RESULT 8
US-09-972-211-128
; Sequence 128, Application US/09972211
; Publication No. US20040048245A1
; GENERAL INFORMATION:
; APPLICANT: Shinkets, Richard A
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Grosse, William M
; APPLICANT: Alsbrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Spytek, Kimberly Ann
; APPLICANT: Li, Li
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Eilerman, Karen
; APPLICANT: MacDougall, John R
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David J
; APPLICANT: Smithson, Glennda
; APPLICANT: Szekeres Jr, Edward S
; TITLE OF INVENTION: No. US20040048245A1el Human Proteins, Polynucleotides Encoding Th
; TITLE OF INVENTION: Methods Of Using The Same
; FILE REFERENCE: 21402-141

```

```

; APPLICANT: Guo, Xiaojia
; APPLICANT: Miller, Charles E.
; APPLICANT: Gangolli, Esha A.
; TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-214 CIP
; CURRENT APPLICATION NUMBER: US/10/087,684
; CURRENT FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: 60/253,834
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,926
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/264,180
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/313,656
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/327,456
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 220
; SOFTWARE: Curaseqlist version 0.1
; SEQ ID NO 87
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Domain
US-10-087-684-87

```

```

Query Match 41.4%; Score 57.5; DB 15; Length 95;
Best Local Similarity 53.8%; Pred. No. 1.2;
Matches 14; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

```

```

Qy 2 ERBERLAYIADH-LGFSWTELARAL 26
Db 9 ELTREKLAKLDHDLGDDWRELARKL 34

```

```

RESULT 10
US-10-218-779-87
; Sequence 87, Application US/10218779
; Publication No. US2004009222A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit
; APPLICANT: MacDougall, John
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William
; APPLICANT: Alsobrook II, John
; APPLICANT: Lepley, Denise
; APPLICANT: Rieger, Daniel
; APPLICANT: Burgess, Catherine
; APPLICANT: Casman, Stacie
; APPLICANT: Spytek, Kimberly
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Murulidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Patturajan, Meera
; APPLICANT: Shenoy, Suresh
; APPLICANT: Rastelli, Luca
; APPLICANT: Tchernev, Velizar
; APPLICANT: Vernet, Corine
; APPLICANT: Zerkhusen, Bryan
; APPLICANT: Malyankar, Uriel
; APPLICANT: Guo, Xiaojia
; APPLICANT: Miller, Charles
; APPLICANT: Gangolli, Esha
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-214
; CURRENT APPLICATION NUMBER: US/10/218,779

```

```

; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: 60/253,834
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,926
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/264,180
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/313,656
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/327,456
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-218-779-87

```

```

Query Match 41.4%; Score 57.5; DB 15; Length 95;
Best Local Similarity 53.8%; Pred. No. 1.2;
Matches 14; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

```

```

Qy 2 ERBERLAYIADH-LGFSWTELARAL 26
Db 9 ELTREKLAKLDHDLGDDWRELARKL 34

```

```

RESULT 11
US-10-096-625-128
; Sequence 128, Application US/10096625
; Publication No. US20040068095A1
; GENERAL INFORMATION:
; APPLICANT: Shinkets, Richard A
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerkhusen, Bryan D
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Spytek, Kimberly Ann
; APPLICANT: Li, Li
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John R
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David J
; APPLICANT: Smithson, Glennnda
; APPLICANT: Szekeres Jr, Edward S
; APPLICANT: Ji, Weizhen
; TITLE OF INVENTION: No. US20040068095A1el Human Proteins, Polynucleotides Encoding The
; FILE REFERENCE: 21402-141 CIP
; CURRENT APPLICATION NUMBER: US/10/096,625
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/972,211
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/238,325
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,323
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,400
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,397
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,401
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,379

```



```

/ GENERAL INFORMATION:
/ APPLICANT: Kekuda, Ramesh
/ APPLICANT: Alsobrook II, John P
/ APPLICANT: Tchernev, Velizar T
/ APPLICANT: Liu, Xiaohong
/ APPLICANT: Spytek, Kimberly A
/ APPLICANT: Patturajan, Meera
/ APPLICANT: Grosse, William M
/ APPLICANT: Lepley, Denise M
/ APPLICANT: Burgess, Catherine E
/ APPLICANT: Vernet, Corine A.M.
/ APPLICANT: Li, Li
/ APPLICANT: Gorman, Linda
/ APPLICANT: Edinger, Shlomit R
/ APPLICANT: Sciore, Paul
/ APPLICANT: Ellerman, Karen
/ APPLICANT: Malyankar, Uriel M
/ APPLICANT: Rothenberg, Mark
/ APPLICANT: Stone, David J
/ APPLICANT: Boldog, Ferenc L
/ APPLICANT: Guo, Xiaojia
/ APPLICANT: Shenoy, Suresh G
/ APPLICANT: Anderson, David W
/ APPLICANT: Padigar, Muralidhara
/ APPLICANT: Taupier Jr, Raymond J
/ APPLICANT: Miller, Charles E
/ APPLICANT: Eisen, Andrew J
/ TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-235
/ CURRENT APPLICATION NUMBER: US/10/037,417
/ CURRENT FILING DATE: 2002-09-20
/ PRIOR APPLICATION NUMBER: 60/260,018
/ PRIOR FILING DATE: 2001-01-05
/ PRIOR APPLICATION NUMBER: 60/260,360
/ PRIOR FILING DATE: 2001-01-08
/ PRIOR APPLICATION NUMBER: 60/272,411
/ PRIOR FILING DATE: 2001-02-28
/ PRIOR APPLICATION NUMBER: 60/272,817
/ PRIOR FILING DATE: 2001-03-02
/ PRIOR APPLICATION NUMBER: 60/291,186
/ PRIOR FILING DATE: 2001-05-15
/ PRIOR APPLICATION NUMBER: 60/303,231
/ PRIOR FILING DATE: 2001-07-05
/ PRIOR APPLICATION NUMBER: 60/305,060
/ PRIOR FILING DATE: 2001-07-12
/ PRIOR APPLICATION NUMBER: 60/318,405
/ PRIOR FILING DATE: 2001-09-10
/ PRIOR APPLICATION NUMBER: 60/318,700
/ PRIOR FILING DATE: 2001-09-12
/ NUMBER OF SEQ ID NOS: 227
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 124
/ LENGTH: 96
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: DEATH domain
/ OTHER INFORMATION: Consensus Sequence
US-10-037-417-124

Query Match 41.4%; Score 57.5; DB 15; Length 96;
Best Local Similarity 53.8%; Pred. No. 1.2;
Matches 14; Conservative 2; Mismatches 9; Indels 1; Gaps

QY 2 ERIEERLXIADH-LGFSWTELARAL 26
DB 9 ELTREKLAKLDDLDGDDWRELARKL 34

RESULT 14
US-10-156-761-9326
/ Sequence 9326, Application US/10156761
/ Publication No. US20030119018A1

```

```

; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9326
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9326

Query Match          40.3%; Score 56; DB 14; Length 384;
Best Local Similarity 48.0%; Pred. No. 9.3;
Matches 12; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

```

```

QY 1 WERIEERLAYIADHGLGFSWTELARA 25
Db 37 WHTTEERAQALDHLGLAGRAEPRA 61

```

```

RESULT 15
US-10-151-208-7
Sequence 7, Application US/10151208
Publication No. US20030105285A1
; GENERAL INFORMATION:
; APPLICANT: Ngai, John
; APPLICANT: Speca, David J.
; APPLICANT: Lin, David M.
; APPLICANT: Isacoff, Ehud Y.
; APPLICANT: Dittman, Andrew H.
; APPLICANT: Fan, Jinhong
; TITLE OF INVENTION: Odorant Receptors
; FILE REFERENCE: B99-038-2
; CURRENT APPLICATION NUMBER: US/10/151,208
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: US/09/619,353
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 60/144,766
; PRIOR FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 7
; LENGTH: 835
; TYPE: PRT
; ORGANISM: Carassius auratus
US-10-151-208-7

Query Match          37.4%; Score 52; DB 14; Length 835;
Best Local Similarity 75.0%; Pred. No. 83;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 8 LAVIADHGLGFSW 19
Db 199 LAVIVKHLGWSW 210

```

Search completed: November 10, 2004, 16:36:05
Job time : 26.0772 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:36:51 ; Search time 5.25617 Seconds
(without alignments)
475.942 Million cell updates/sec

Title: US-10-092-750-21

Perfect score: 139

Sequence: 1 WERIERLAYIADHLGFSWTELARAL 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 123 | 88.5 | 3924 | 2 S37431 | ankyrin 2, neuropa |
| 2 | 90 | 64.7 | 1765 | 2 T42714 | ankyrin 3, splice |
| 3 | 90 | 64.7 | 1940 | 2 T42715 | ankyrin 3, splice |
| 4 | 90 | 64.7 | 1943 | 2 T42713 | ankyrin 3, splice |
| 5 | 90 | 64.7 | 1961 | 2 T42716 | ankyrin 3, splice |
| 6 | 90 | 64.7 | 4377 | 2 A55575 | ankyrin 3, long sp |
| 7 | 82 | 59.0 | 138 | 2 S37773 | ankyrin, erythrocy |
| 8 | 82 | 59.0 | 1848 | 2 S37771 | ankyrin, erythrocy |
| 9 | 82 | 59.0 | 1862 | 2 I49502 | ankyrin - mouse |
| 10 | 76 | 54.7 | 1856 | 2 B35049 | ankyrin 1, erythro |
| 11 | 76 | 54.7 | 1880 | 2 A35049 | ankyrin 1, erythro |
| 12 | 76 | 54.7 | 1881 | 1 SUHUK | ankyrin 1, erythro |
| 13 | 57 | 41.0 | 311 | 2 F69820 | conserved hypother |
| 14 | 52 | 37.4 | 489 | 2 T36100 | probable ATP-bind |
| 15 | 52 | 37.4 | 496 | 2 C64935 | hypothetical prote |
| 16 | 52 | 37.4 | 496 | 2 E90936 | probable transport |
| 17 | 52 | 37.4 | 496 | 2 A85785 | probable transport |
| 18 | 51 | 36.7 | 484 | 2 B64481 | hypothetical prote |
| 19 | 51 | 36.7 | 613 | 2 T00077 | gag-like protein - |
| 20 | 50.5 | 36.3 | 757 | 1 I50667 | transforming prote |
| 21 | 50 | 36.0 | 416 | 2 S70090 | hypothetical prote |
| 22 | 49.5 | 35.6 | 258 | 2 T36359 | hypothetical prote |
| 23 | 49.5 | 35.6 | 348 | 2 S11198 | transforming prote |
| 24 | 49.5 | 35.6 | 388 | 1 COIV | transforming prote |
| 25 | 49.5 | 35.6 | 636 | 1 TVMSMB | transforming prote |
| 26 | 49.5 | 35.6 | 640 | 1 A55073 | transforming prote |
| 27 | 49.5 | 35.6 | 715 | 4 TVMSMY | transforming prote |
| 28 | 49.5 | 35.6 | 751 | 1 I49497 | transforming prote |
| 29 | 49.5 | 35.6 | 752 | 1 S03423 | transforming prote |

ALIGNMENTS

RESULT 1

S37431

ankyrin 2, neuronal long splice form - human

N:Alternate names: ankyrin B, 440K splice form; ankyrin-B; brain ankyrin; non-erythroid

N:Contains: ankyrin 2, short form

C:Species: Homo sapiens (man)

C>Date: 06-Jan-1995 #sequence_revision 05-Jan-1995 #text change 09-Jul-2004

C:Accession: S37431; A39643; B39643; A40334; A49462; S14533; S14569

R:Chan, W.

submitted to the EMBL Data Library, September 1993

A:Reference number: S37431

A:Accession: S37431

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-3924 <CHA>

A:Cross-references: UNIPROT:Q01484; EMBL:Z26634; NID:G406287; PIDN:CAA81387.1; PID:G40628

R:Otsu, E.; Kunitomo, M.; McLaughlin, T.; Bennett, V.

J. Cell Biol. 114, 241-253, 1991

A:Title: Isolation and characterization of cDNAs encoding human brain ankyrins reveal a

A:Reference number: A39643; MUID:91302466; PMID:1830053

A:Accession: A39643

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-2077 <OT1>

A:Cross-references: GB:X56957

A:Accession: B39643

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1443,3585-3924 <OTT>

A:Cross-references: EMBL:X56958

R:Tse, W.T.; Menninger, J.C.; Yang-Feng, T.L.; Francke, U.; Sahr, K.E.; Lux, S.E.; Ward,

Genomics 10, 858-866, 1991

A:Title: Isolation and chromosomal localization of a novel nonerythroid ankyrin gene.

A:Reference number: A40334; MUID:92009921; PMID:1833308

A:Accession: A40334

A:Molecule type: DNA

A:Residues: 463-474, 'PE', 477-495 <TSE>

A:Cross-references: GB:M37123; NID:G178647; PIDN:AAA62828.1; PID:G178648

R:Chan, W.; Kordeli, E.; Bennett, V.

J. Cell Biol. 123, 1463-1473, 1993

A:Title: 440-kD ankyrinB: structure of the major developmentally regulated domain and sel

A:Reference number: A49462; MUID:94075409; PMID:8253844

A:Accession: A49462

A>Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-3924 <RES>

A:Cross-references: EMBL:Z26634; NID:G406287; PIDN:CAA81387.1; PID:G406288

C:Genetics:

A:Gene: GDB:ANK2

A:Cross-references: GDB:127607; OMIM:106410

A:Map position: 4q25-4q27

C:Superfamily: ankyrin; ankyrin repeat homology

C;Keywords: alternative splicing
 F;2-3924/Product: ankryrin 2, long form #status predicted <MAT>
 F;2-1443,3585-3924/Product: ankryrin 2, short form #status predicted <MA2>
 F;63-98/Domain: ankryrin repeat homology <AN01>
 F;96-128/Domain: ankryrin repeat homology <AN02>
 F;129-161/Domain: ankryrin repeat homology <AN03>
 F;162-190/Domain: ankryrin repeat homology <AN04>
 F;191-223/Domain: ankryrin repeat homology <AN05>
 F;232-264/Domain: ankryrin repeat homology <AN06>
 F;265-297/Domain: ankryrin repeat homology <AN07>
 F;298-330/Domain: ankryrin repeat homology <AN08>
 F;331-363/Domain: ankryrin repeat homology <AN09>
 F;364-396/Domain: ankryrin repeat homology <AN10>
 F;397-429/Domain: ankryrin repeat homology <AN11>
 F;430-462/Domain: ankryrin repeat homology <AN12>
 F;463-495/Domain: ankryrin repeat homology <AN13>
 F;496-528/Domain: ankryrin repeat homology <AN14>
 F;529-561/Domain: ankryrin repeat homology <AN15>
 F;562-594/Domain: ankryrin repeat homology <AN16>
 F;595-627/Domain: ankryrin repeat homology <AN17>
 F;628-660/Domain: ankryrin repeat homology <AN18>
 F;661-693/Domain: ankryrin repeat homology <AN19>
 F;694-726/Domain: ankryrin repeat homology <AN20>
 F;727-759/Domain: ankryrin repeat homology <AN21>
 F;760-792/Domain: ankryrin repeat homology <AN22>
 F;793-825/Domain: ankryrin repeat homology <AN23>
 Query Match 88.5%; Score 123; DB 2; Length 3924;
 Best Local Similarity 96.0%; Pred. No. 8.8e-09;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 ERIBERLAYIADHLGFSWTELARAL 26
 Db 3534 ERIBERLAYIADHLGFSWTELARAL 3558
 RESULT 2
 T42714
 ankryrin 3, splice form 2 - mouse
 C;Species: Mus musculus (house mouse)
 C;Accession: T42714
 R;Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, M.; Turtzo, I.
 J. Cell Biol. 130, 313-330, 1995
 A;Title: Ank3 (epithelial ankryrin), a widely distributed new member of the ankryrin gene
 the repeat domain.
 A;Reference number: 22237; MUID:95340633; PMID:7615634
 A;Accession: T42714
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-1765 <PET>
 A;Cross-references: UNIPROT:Q61307; EMBL:L40632; NID:g710548; PID:g710551; PIDN:AAB01605
 A;Experimental source: strain C57BL/6J; kidney
 C;Genetics:
 A;Gene: Ank3
 A;Map position: 10
 A;Introns: 1587/1
 C;Superfamily: ankryrin; ankryrin repeat homology
 C;Keywords: alternative splicing
 Query Match 64.7%; Score 90; DB 2; Length 1765;
 Best Local Similarity 68.0%; Pred. No. 0.00019;
 Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 2 ERIBERLAYIADHLGFSWTELARAL 26
 Db 1476 ERIBERLAYIADHLGFSWTELARAL 1500
 RESULT 3
 T42715
 ankryrin 3, splice form 3 - mouse
 C;Species: Mus musculus (house mouse)

C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
 C;Accession: T42715
 R;Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, L.
 J. Cell Biol. 130, 313-330, 1995
 A;Title: Ank3 (epithelial ankryrin), a widely distributed new member of the ankryrin gene
 the repeat domain.
 A;Reference number: 22237; MUID:95340633; PMID:7615634
 A;Accession: T42715
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-1940 <PET>
 A;Cross-references: UNIPROT:Q61307; EMBL:L40632; NID:g710548; PID:g710549; PIDN:AAB01604
 A;Experimental source: strain C57BL/6J; kidney
 C;Genetics:
 A;Gene: Ank3
 A;Map position: 10
 A;Introns: 834/1
 C;Superfamily: ankryrin; ankryrin repeat homology
 C;Keywords: alternative splicing
 Query Match 64.7%; Score 90; DB 2; Length 1940;
 Best Local Similarity 68.0%; Pred. No. 0.00021;
 Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 2 ERIBERLAYIADHLGFSWTELARAL 26
 Db 1455 ERIBERLAYIADHLGFSWTELARAL 1479
 RESULT 4
 T42713
 ankryrin 3, splice form 1 - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
 C;Accession: T42713
 R;Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, L.
 J. Cell Biol. 130, 313-330, 1995
 A;Title: Ank3 (epithelial ankryrin), a widely distributed new member of the ankryrin gene
 the repeat domain.
 A;Reference number: 22237; MUID:95340633; PMID:7615634
 A;Accession: T42713
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-1943 <PET>
 A;Cross-references: UNIPROT:Q61307; EMBL:L40632; NID:g710548; PID:g710550; PIDN:AAB01606
 A;Experimental source: strain C57BL/6J; kidney
 C;Genetics:
 A;Gene: Ank3
 A;Map position: 10
 A;Introns: 855/1
 C;Function:
 A;Description: supposed to play an important role in the polarized distribution of many
 A;Note: major kidney ankryrin
 C;Superfamily: ankryrin; ankryrin repeat homology
 C;Keywords: alternative splicing
 Query Match 64.7%; Score 90; DB 2; Length 1943;
 Best Local Similarity 68.0%; Pred. No. 0.00021;
 Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 2 ERIBERLAYIADHLGFSWTELARAL 26
 Db 1458 ERIBERLAYIADHLGFSWTELARAL 1482
 RESULT 5
 T42716
 ankryrin 3, splice form 4 - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
 C;Accession: T42716
 R;Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, L.
 J. Cell Biol. 130, 313-330, 1995

| | | | | |
|-----------------------|--------------|------------------|--------------|-------------|
| Query Match | 64.7% | Score 90 | DB 2 | Length 4377 |
| Best Local Similarity | 68.0% | Pred. No. 0.0005 | | |
| Matches 17 | Conservative | 3 | Mismatches 5 | Indels 0 |
| | | | | Gaps 0 |

F;638-670/Domain: ankyrin repeat homology <A
F;671-703/Domain: ankyrin repeat homology <A
F;704-736/Domain: ankyrin repeat homology <A

Query Match 59.0%; Score 82; DB 2; Length 1848;

```

A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1856 <LAM>
C:Genetics:
A:Gene: GDB:ANK1; ANK
A:Cross-references: GDB:118737; OMIM:182900
A:Map position: 8p11.2-8p11.2
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing
F:2-1856/Product: ankyrin 1, erythrocyte form 3 #status predicted <MAT>
F:2-1513.1676-1856/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>
F:44-76/Domain: ankyrin repeat homology <AN01>
F:77-109/Domain: ankyrin repeat homology <AN02>
F:110-142/Domain: ankyrin repeat homology <AN03>
F:143-171/Domain: ankyrin repeat homology <AN04>
F:172-204/Domain: ankyrin repeat homology <AN05>
F:205-237/Domain: ankyrin repeat homology <AN06>
F:238-270/Domain: ankyrin repeat homology <AN07>
F:271-303/Domain: ankyrin repeat homology <AN08>
F:304-336/Domain: ankyrin repeat homology <AN09>
F:337-369/Domain: ankyrin repeat homology <AN10>
F:370-402/Domain: ankyrin repeat homology <AN11>
F:403-435/Domain: ankyrin repeat homology <AN12>
F:436-468/Domain: ankyrin repeat homology <AN13>
F:469-501/Domain: ankyrin repeat homology <AN14>
F:502-534/Domain: ankyrin repeat homology <AN15>
F:535-567/Domain: ankyrin repeat homology <AN16>
F:568-600/Domain: ankyrin repeat homology <AN17>
F:601-633/Domain: ankyrin repeat homology <AN18>
F:634-666/Domain: ankyrin repeat homology <AN19>
F:667-699/Domain: ankyrin repeat homology <AN20>
F:700-732/Domain: ankyrin repeat homology <AN21>
F:733-765/Domain: ankyrin repeat homology <AN22>
F:766-798/Domain: ankyrin repeat homology <AN23>

Query Match 54.7%; Score 76; DB 2; Length 1856;
Best Local Similarity 56.0%; Pred. NO. 0.02;
Matches 14; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 ERIERLAVIADHLGFSWTELARAL 26
DB 1401 EQAEKMKVAISEHLGLSWAELAREL 1425
| : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : :

RESULT 11
A35049
ankyrin 1, erythrocyte splice form 2 - human
N:Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
N:Contains: ankyrin 2.2, erythrocyte
C:Species: Homo sapiens (man)
C>Date: 27-Jul-1990 #sequence_revision 01-Oct-1992 #text_change 09-Jul-2004
C:Accession: A35049
R:Labmet, S.; Yu, H.; Pritch, J.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.; Krieger, N. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A:Title: cDNA sequence for human erythrocyte ankyrin.
A:Reference number: A35049; MUID:90175370; PMID:1689849
A:Accession: A35049
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1880 <LAM>
A:Cross-references: UNIPROT:P16157; GB:M28880
C:Genetics:
A:Gene: GDB:ANK1; ANK
A:Cross-references: GDB:118737; OMIM:182900
A:Map position: 8p11.2-8p11.2
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing; cytoskeleton
F:2-1880/Product: ankyrin 1, erythrocyte form 2 #status predicted <MAT>
F:2-1513.1676-1880/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>
F:44-76/Domain: ankyrin repeat homology <AN01>
F:77-109/Domain: ankyrin repeat homology <AN02>
F:110-142/Domain: ankyrin repeat homology <AN03>
F:143-171/Domain: ankyrin repeat homology <AN04>

```

F;2-827/Region: anion exchange protein binding
F;44-76/Domain: ankryrin repeat homology <AN01>
F;77-109/Domain: ankryrin repeat homology <AN02>
F;110-142/Domain: ankryrin repeat homology <AN03>
F;143-171/Domain: ankryrin repeat homology <AN04>
F;172-204/Domain: ankryrin repeat homology <AN05>
F;205-237/Domain: ankryrin repeat homology <AN06>
F;238-270/Domain: ankryrin repeat homology <AN07>
F;271-303/Domain: ankryrin repeat homology <AN08>
F;304-336/Domain: ankryrin repeat homology <AN09>
F;337-369/Domain: ankryrin repeat homology <AN10>
F;370-402/Domain: ankryrin repeat homology <AN11>
F;403-435/Domain: ankryrin repeat homology <AN12>
F;436-468/Domain: ankryrin repeat homology <AN13>
F;469-501/Domain: ankryrin repeat homology <AN14>
F;502-534/Domain: ankryrin repeat homology <AN15>
F;535-567/Domain: ankryrin repeat homology <AN16>
F;568-600/Domain: ankryrin repeat homology <AN17>
F;601-633/Domain: ankryrin repeat homology <AN18>
F;634-666/Domain: ankryrin repeat homology <AN19>
F;667-699/Domain: ankryrin repeat homology <AN20>
F;700-732/Domain: ankryrin repeat homology <AN21>
F;733-765/Domain: ankryrin repeat homology <AN22>
F;766-798/Domain: ankryrin repeat homology <AN23>
F;828-1382/Domain: 65K #status predicted <DOM2>
F;828-1382/Region: spectrin binding
F;1383-1981/Domain: 55K #status predicted <DOM3>

Query Match 54.7%; Score 76; DB 1; Length 1881;
Best Local Similarity 56.0%; Pred. No. 0.02;
Matches 14; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 ERTEERLAYIADHLGFSWTELARAL 26
| : | : | : | : | : | : | : | : | : |
DB 1401 EQAEMKNVAIVSEHGLGSWAELAREL 1425

RESULT 13
F69820
conserved hypothetical protein yhbB - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: F69820
R;Kunst, S.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero,
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi,
A.; Ehrlich, S.D.; Emmerson, P.T.; Enrian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallercia,
J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hulio, M.F.;
Koster, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
M.; Ogawa, K.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel,
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Ross, M.; Sadie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scrofene, F.; Sekiguchi, J.; Sekowska, A.; Seror,
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yanane, K.; Yasumoto, K.; Yata, K.; Yoshikawa, K.;
A;Authors: Yoshikawa, H.F.; Zumanst, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: F69820
A;Status: Preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-311 <KUN>
A;Cross-references: UNIPROT:O31589; GB:Z39108; GB:AL009126; NID:G2633055; PIDN:CAB12720.1
A;Experimental source: strain 168
C;Genetics:
A;Gene: yhbB

Query Match 41.0%; Score 57; DB 2; Length 311;
Best Local Similarity 47.8%; Pred. No. 1.6;
Matches 11; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 2 ERTEERLAYIADHLGFSWTELAR 24

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 11:27:01 ; Search time 28.4475 Seconds
(without alignments)
525.871 Million cell updates/sec

Title: US-10-092-750-21

Perfect score: 139

Sequence: 1 WRIERLAYIADHLGFSWTELARAL 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt C2:*

1: uniprot_prot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|--------------|---------------------|
| 1 | 123 | 88.5 | 965 | 2 Q72344 | Q72344 homo sapien |
| 2 | 123 | 88.5 | 1863 | 2 Q723L5 | Q723L5 homo sapien |
| 3 | 123 | 88.5 | 3924 | 1 ANK2_HUMAN | Q01484 homo sapien |
| 4 | 120 | 86.3 | 1038 | 2 Q80ZZ7 | Q80ZZ7 mus musculus |
| 5 | 120 | 86.3 | 1050 | 2 Q6PCN2 | Q6PCN2 mus musculus |
| 6 | 120 | 86.3 | 1050 | 2 Q8CCV0 | Q8CCV0 mus musculus |
| 7 | 120 | 86.3 | 1050 | 2 AAH59251 | AAH59251 mus muscu |
| 8 | 113 | 81.3 | 202 | 2 Q9BEY7 | Q9BEY7 poephila gu |
| 9 | 90 | 64.7 | 143 | 2 Q70510 | Q70510 rattus norv |
| 10 | 90 | 64.7 | 811 | 2 Q9H0P5 | Q9H0P5 homo sapien |
| 11 | 90 | 64.7 | 838 | 2 Q9QXH1 | Q9QXH1 mus musculus |
| 12 | 90 | 64.7 | 960 | 2 Q8VDA0 | Q8VDA0 rattus norv |
| 13 | 90 | 64.7 | 1093 | 2 Q8CBN3 | Q8CBN3 mus musculus |
| 14 | 90 | 64.7 | 1101 | 2 Q62T73 | Q62T73 homo sapien |
| 15 | 90 | 64.7 | 1101 | 2 BAC86721 | BAC86721 homo sapi |
| 16 | 90 | 64.7 | 1726 | 2 Q8VC68 | Q8VC68 mus musculus |
| 17 | 90 | 64.7 | 1762 | 2 Q88521 | Q88521 rattus norv |
| 18 | 90 | 64.7 | 1887 | 2 Q723G4 | Q723G4 homo sapien |
| 19 | 90 | 64.7 | 1943 | 2 Q61307 | Q61307 mus musculus |
| 20 | 90 | 64.7 | 2622 | 2 Q70511 | Q70511 rattus norv |
| 21 | 90 | 64.7 | 4377 | 1 ANK3_HUMAN | Q12955 homo sapien |
| 22 | 82 | 59.0 | 1098 | 2 Q61304 | Q61304 mus musculus |
| 23 | 82 | 59.0 | 1848 | 2 Q61302 | Q61302 mus musculus |
| 24 | 82 | 59.0 | 1862 | 1 ANK1_MOUSE | Q02357 mus musculus |
| 25 | 76 | 54.7 | 1719 | 2 Q13768 | Q13768 homo sapien |
| 26 | 76 | 54.7 | 1856 | 2 Q99407 | Q99407 homo sapien |
| 27 | 76 | 54.7 | 1880 | 1 ANK1_HUMAN | P16157 homo sapien |
| 28 | 60 | 43.2 | 1021 | 2 Q758Z9 | Q758Z9 ashbysa goss |
| 29 | 60 | 43.2 | 1021 | 2 AA52298 | AA52298 ashbysa go |
| 30 | 57 | 41.0 | 311 | 2 P97031 | P97031 bacillus su |
| 31 | 57 | 41.0 | 311 | 2 O31589 | O31589 bacillus su |

32 56 40.3 384 2 Q82M73 Q82M73 streptomyc
33 55 39.6 465 2 Q8EPA9 Q8EPA9 oceanobacil
34 54 38.8 433 2 Q72K95 Q72K95 thermus t
35 54 38.8 433 2 AAS8Q910 AAS8Q910 thermus t
36 54 38.8 605 2 Q9P5X0 Q9P5X0 neurospora
37 53.5 38.5 1201 2 Q6CE33 Q6CE33 varrowia li
38 53 38.1 77 2 Q8E4N4 Q8E4N4 streptococc
39 52.5 37.8 602 2 Q75NZ5 Q75NZ5 chlamydomon
40 52.5 37.8 602 2 BAD13491 BAD13491 chlamydom
41 52.5 37.8 602 2 BAD13492 BAD13492 chlamydom
42 52 37.4 57 2 Q87KE7 Q87KE7 vibrio para
43 52 37.4 157 2 Q8SUB7 Q8SUB7 encephalito
44 52 37.4 215 2 Q7QVM4 Q7QVM4 giardia lam
45 52 37.4 335 2 Q725E8 Q725E8 desulfovibr

ALIGNMENTS

RESULT 1
Q72344 PRELIMINARY: PRT; 965 AA.
AC Q72344
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Hypothetical protein DkFZp686M09125 (Fragment).
GN Name=DkFZp686M09125;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=Human uterus;
RA Bloeker H., Boscher M., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX538132; CAD98033.1; -
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000906; ZUS.
DR Pfam; PF00531; Death; 1.
DR Pfam; PF00791; ZUS; 1.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZUS; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 965 AA; 108618 MW; FCC332E58505EE43 CRC64;
Query Match 88.5%; Score 123; DB 2; Length 965;
Best Local Similarity 96.0%; Pred.No.8.le-09;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 ERIERLAYIADHLGFSWTELARAL 26
Db 483 ERIERLAYIADHLGFSWTELAREL 507
RESULT 2
Q723L5 PRELIMINARY: PRT; 1863 AA.
AC Q723L5
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Hypothetical protein DkFZp686H0688.
GN Name=DkFZp686H0688;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human retina;
RA Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX537758; CAD97827.1; -;
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000906; ZUS.
DR Pfam; PF00023; Ank; 23.
DR Pfam; PF00531; Death; 1.
DR Pfam; PF00731; ZUS; 1.
DR SMART; SM00248; ANK; 23.
DR SMART; SM00005; DEATH; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW ANK repeat: Hypothetical protein.
SQ SEQUENCE 1863 AA; 204736 MW; 184C998E0F0A03DF CRC64;

Query Match 88.5%; Score 123; DB 2; Length 1863;
Best Local Similarity 96.0%; Pred. No. 1.6e-08;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ERIERLAYIADHLGFSWTELARAL 26
Db 1473 ERIERLAYIADHLGFSWTELAREL 1497
[1]
[2]
[3]
[4]
[5]
[6]
[7]
[8]
[9]
[10]
[11]
[12]
[13]
[14]
[15]
[16]
[17]
[18]
[19]
[20]
[21]
[22]
[23]
[24]
[25]
[26]
[27]
[28]
[29]
[30]
[31]
[32]
[33]
[34]
[35]
[36]
[37]
[38]
[39]
[40]
[41]
[42]
[43]
[44]
[45]
[46]
[47]
[48]
[49]
[50]
[51]
[52]
[53]
[54]
[55]
[56]
[57]
[58]
[59]
[60]
[61]
[62]
[63]
[64]
[65]
[66]
[67]
[68]
[69]
[70]
[71]
[72]
[73]
[74]
[75]
[76]
[77]
[78]
[79]
[80]
[81]
[82]
[83]
[84]
[85]
[86]
[87]
[88]
[89]
[90]
[91]
[92]
[93]
[94]
[95]
[96]
[97]
[98]
[99]
[100]
[101]
[102]
[103]
[104]
[105]
[106]
[107]
[108]
[109]
[110]
[111]
[112]
[113]
[114]
[115]
[116]
[117]
[118]
[119]
[120]
[121]
[122]
[123]
[124]
[125]
[126]
[127]
[128]
[129]
[130]
[131]
[132]
[133]
[134]
[135]
[136]
[137]
[138]
[139]
[140]
[141]
[142]
[143]
[144]
[145]
[146]
[147]
[148]
[149]
[150]
[151]
[152]
[153]
[154]
[155]
[156]
[157]
[158]
[159]
[160]
[161]
[162]
[163]
[164]
[165]
[166]
[167]
[168]
[169]
[170]
[171]
[172]
[173]
[174]
[175]
[176]
[177]
[178]
[179]
[180]
[181]
[182]
[183]
[184]
[185]
[186]
[187]
[188]
[189]
[190]
[191]
[192]
[193]
[194]
[195]
[196]
[197]
[198]
[199]
[200]
[201]
[202]
[203]
[204]
[205]
[206]
[207]
[208]
[209]
[210]
[211]
[212]
[213]
[214]
[215]
[216]
[217]
[218]
[219]
[220]
[221]
[222]
[223]
[224]
[225]
[226]
[227]
[228]
[229]
[230]
[231]
[232]
[233]
[234]
[235]
[236]
[237]
[238]
[239]
[240]
[241]
[242]
[243]
[244]
[245]
[246]
[247]
[248]
[249]
[250]
[251]
[252]
[253]
[254]
[255]
[256]
[257]
[258]
[259]
[260]
[261]
[262]
[263]
[264]
[265]
[266]
[267]
[268]
[269]
[270]
[271]
[272]
[273]
[274]
[275]
[276]
[277]
[278]
[279]
[280]
[281]
[282]
[283]
[284]
[285]
[286]
[287]
[288]
[289]
[290]
[291]
[292]
[293]
[294]
[295]
[296]
[297]
[298]
[299]
[300]
[301]
[302]
[303]
[304]
[305]
[306]
[307]
[308]
[309]
[310]
[311]
[312]
[313]
[314]
[315]
[316]
[317]
[318]
[319]
[320]
[321]
[322]
[323]
[324]
[325]
[326]
[327]
[328]
[329]
[330]
[331]
[332]
[333]
[334]
[335]
[336]
[337]
[338]
[339]
[340]
[341]
[342]
[343]
[344]
[345]
[346]
[347]
[348]
[349]
[350]
[351]
[352]
[353]
[354]
[355]
[356]
[357]
[358]
[359]
[360]
[361]
[362]
[363]
[364]
[365]
[366]
[367]
[368]
[369]
[370]
[371]
[372]
[373]
[374]
[375]
[376]
[377]
[378]
[379]
[380]
[381]
[382]
[383]
[384]
[385]
[386]
[387]
[388]
[389]
[390]
[391]
[392]
[393]
[394]
[395]
[396]
[397]
[398]
[399]
[400]
[401]
[402]
[403]
[404]
[405]
[406]
[407]
[408]
[409]
[410]
[411]
[412]
[413]
[414]
[415]
[416]
[417]
[418]
[419]
[420]
[421]
[422]
[423]
[424]
[425]
[426]
[427]
[428]
[429]
[430]
[431]
[432]
[433]
[434]
[435]
[436]
[437]
[438]
[439]
[440]
[441]
[442]
[443]
[444]
[445]
[446]
[447]
[448]
[449]
[450]
[451]
[452]
[453]
[454]
[455]
[456]
[457]
[458]
[459]
[460]
[461]
[462]
[463]
[464]
[465]
[466]
[467]
[468]
[469]
[470]
[471]
[472]
[473]
[474]
[475]
[476]
[477]
[478]
[479]
[480]
[481]
[482]
[483]
[484]
[485]
[486]
[487]
[488]
[489]
[490]
[491]
[492]
[493]
[494]
[495]
[496]
[497]
[498]
[499]
[500]
[501]
[502]
[503]
[504]
[505]
[506]
[507]
[508]
[509]
[510]
[511]
[512]
[513]
[514]
[515]
[516]
[517]
[518]
[519]
[520]
[521]
[522]
[523]
[524]
[525]
[526]
[527]
[528]
[529]
[530]
[531]
[532]
[533]
[534]
[535]
[536]
[537]
[538]
[539]
[540]
[541]
[542]
[543]
[544]
[545]
[546]
[547]
[548]
[549]
[550]
[551]
[552]
[553]
[554]
[555]
[556]
[557]
[558]
[559]
[560]
[561]
[562]
[563]
[564]
[565]
[566]
[567]
[568]
[569]
[570]
[571]
[572]
[573]
[574]
[575]
[576]
[577]
[578]
[579]
[580]
[581]
[582]
[583]
[584]
[585]
[586]
[587]
[588]
[589]
[590]
[591]
[592]
[593]
[594]
[595]
[596]
[597]
[598]
[599]
[600]
[601]
[602]
[603]
[604]
[605]
[606]
[607]
[608]
[609]
[610]
[611]
[612]
[613]
[614]
[615]
[616]
[617]
[618]
[619]
[620]
[621]
[622]
[623]
[624]
[625]
[626]
[627]
[628]
[629]
[630]
[631]
[632]
[633]
[634]
[635]
[636]
[637]
[638]
[639]
[640]
[641]
[642]
[643]
[644]
[645]
[646]
[647]
[648]
[649]
[650]
[651]
[652]
[653]
[654]
[655]
[656]
[657]
[658]
[659]
[660]
[661]
[662]
[663]
[664]
[665]
[666]
[667]
[668]
[669]
[670]
[671]
[672]
[673]
[674]
[675]
[676]
[677]
[678]
[679]
[680]
[681]
[682]
[683]
[684]
[685]
[686]
[687]
[688]
[689]
[690]
[691]
[692]
[693]
[694]
[695]
[696]
[697]
[698]
[699]
[700]
[701]
[702]
[703]
[704]
[705]
[706]
[707]
[708]
[709]
[710]
[711]
[712]
[713]
[714]
[715]
[716]
[717]
[718]
[719]
[720]
[721]
[722]
[723]
[724]
[725]
[726]
[727]
[728]
[729]
[730]
[731]
[732]
[733]
[734]
[735]
[736]
[737]
[738]
[739]
[740]
[741]
[742]
[743]
[744]
[745]
[746]
[747]
[748]
[749]
[750]
[751]
[752]
[753]
[754]
[755]
[756]
[757]
[758]
[759]
[760]
[761]
[762]
[763]
[764]
[765]
[766]
[767]
[768]
[769]
[770]
[771]
[772]
[773]
[774]
[775]
[776]
[777]
[778]
[779]
[780]
[781]
[782]
[783]
[784]
[785]
[786]
[787]
[788]
[789]
[790]
[791]
[792]
[793]
[794]
[795]
[796]
[797]
[798]
[799]
[800]
[801]
[802]
[803]
[804]
[805]
[806]
[807]
[808]
[809]
[810]
[811]
[812]
[813]
[814]
[815]
[816]
[817]
[818]
[819]
[820]
[821]
[822]
[823]
[824]
[825]
[826]
[827]
[828]
[829]
[830]
[831]
[832]
[833]
[834]
[835]
[836]
[837]
[838]
[839]
[840]
[841]
[842]
[843]
[844]
[845]
[846]
[847]
[848]
[849]
[850]
[851]
[852]
[853]
[854]
[855]
[856]
[857]
[858]
[859]
[860]
[861]
[862]
[863]
[864]
[865]
[866]
[867]
[868]
[869]
[870]
[871]
[872]
[873]
[874]
[875]
[876]
[877]
[878]
[879]
[880]
[881]
[882]
[883]
[884]
[885]
[886]
[887]
[888]
[889]
[890]
[891]
[892]
[893]
[894]
[895]
[896]
[897]
[898]
[899]
[900]
[901]
[902]
[903]
[904]
[905]
[906]
[907]
[908]
[909]
[910]
[911]
[912]
[913]
[914]
[915]
[916]
[917]
[918]
[919]
[920]
[921]
[922]
[923]
[924]
[925]
[926]
[927]
[928]
[929]
[930]
[931]
[932]
[933]
[934]
[935]
[936]
[937]
[938]
[939]
[940]
[941]
[942]
[943]
[944]
[945]
[946]
[947]
[948]
[949]
[950]
[951]
[952]
[953]
[954]
[955]
[956]
[957]
[958]
[959]
[960]
[961]
[962]
[963]
[964]
[965]
[966]
[967]
[968]
[969]
[970]
[971]
[972]
[973]
[974]
[975]
[976]
[977]
[978]
[979]
[980]
[981]
[982]
[983]
[984]
[985]
[986]
[987]
[988]
[989]
[990]
[991]
[992]
[993]
[994]
[995]
[996]
[997]
[998]
[999]
[1000]

```

```
FT DOMAIN 1773 1950 Repeat-rich region.
FT REPEAT 1773 1784 Repeat A.
FT REPEAT 1785 1796 Repeat A.
FT REPEAT 1797 1808 Repeat A.
FT REPEAT 1809 1820 Repeat A.
FT REPEAT 1821 1832 Repeat A.
FT REPEAT 1833 1844 Repeat A.
FT REPEAT 1845 1856 Repeat A.
FT REPEAT 1857 1867 Repeat A (approximate).
FT REPEAT 1868 1879 Repeat A.
FT REPEAT 1880 1891 Repeat A.
FT REPEAT 1892 1902 Repeat A (approximate).
FT REPEAT 1903 1914 Repeat A.
FT REPEAT 1915 1926 Repeat A.
FT REPEAT 1927 1938 Repeat A.
FT REPEAT 1939 1950 Repeat A.
FT DOMAIN 3536 3620 Death.
FT VARSPLOC 1039 1039
Q -> QFLGKLHLPAPPLNEGESIVSRILQLGPPGTK
(in isoform 2).
/FTId=VSP_000267.
Missing (in isoform 2 and isoform 3).
/FTId=VSP_000269.
GQ -> PE (in Ref. 4).
I -> S (in Ref. 1).
QY -> HA (in Ref. 1).
I -> Y (in Ref. 1).
SQ SEQUENCE 3924 AA; 430337 MW; 52AC496C428E29D2 CRC64;

Query Match 88.5%; Score 123; DB 1; Length 3924;
Best Local Similarity 96.0%; Pred. NO. 3.3e-08;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ERIEERLAYIADHLGFSWTELARAL 26
|||||
Db 3534 ERIEERLAYIADHLGFSWTELAREL 3558

RESULT 4
Q80ZZ7 PRELIMINARY; PRT; 1038 AA.
AC Q80ZZ7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to ankryrin 2, neuronal (fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC043123; AAH43123.1; -.
DR GO; GO:0007165; P signal transduction; IEA.
DR InterPro; IPR000488; Death.
DR InterPro; IPR011029; DEATH_like.
DR Pfam; PF00531; Death; 1.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZUS; 1.
DR SMART; SM00005; DEATH; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
SQ SEQUENCE 1038 AA; 114518 MW; 3012ED0B9AA2A0F4 CRC64;

Query Match 86.3%; Score 120; DB 2; Length 1038;
Best Local Similarity 92.0%; Pred. NO. 2.4e-08;
Matches 23; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ERIEERLAYIADHLGFSWTELARAL 26
|||||
Db 599 ERMEERLAYIADHLGFSWTELAREL 623

Query Match 86.3%; Score 120; DB 2; Length 1038;
Best Local Similarity 92.0%; Pred. NO. 2.4e-08;
Matches 23; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ERIEERLAYIADHLGFSWTELARAL 26
|||||
Db 611 ERMEERLAYIADHLGFSWTELAREL 635
```

```
RESULT 5
Q6PCN2 PRELIMINARY; PRT; 1050 AA.
AC Q6PCN2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Ankryrin 2, brain.
GN Name=Ank2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22386257; PubMed=12477932;
RA Strausberg R.L.; Feingold E.A.; Grouse L.H.; Derge J.G.;
RA Klausner R.D.; Collins F.S.; Wagner L.; Shenmen C.M.; Schuler G.D.;
RA Altschul S.F.; Zeeberg B.; Buetow K.H.; Schaefer C.F.; Bhat N.K.;
RA Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Hsieh F.;
RA Diatchenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong J.;
RA Stapleton M.; Soares M.B.; Bonaldo M.F.; Casavant T.D.; Scheetz T.E.;
RA Brownstein M.J.; Usdin T.B.; Toshiyuki S.; Carninci P.; Prange C.;
RA Raha S.S.; Loquellano N.A.; Peters G.J.; Abramson R.D.; Mullahy S.J.;
RA Boak S.A.; McEwan P.J.; McKernan K.J.; Malek J.A.; Gunaratne P.H.;
RA Richards S.; Worley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Hulyk S.W.;
RA Villalon D.K.; Muzny D.M.; Sodergren E.J.; Lu X.; Gibbs R.A.;
RA Faney J.; Helton E.; Kettelman M.; Madan A.; Rodrigues S.; Sanchez A.;
RA Whitting M.; Madan A.; Young A.C.; Shevchenko Y.; Bouffard G.G.;
RA Blakesley R.W.; Touchman J.W.; Green E.D.; Dickson M.C.;
RA Rodriguez A.C.; Grimwood J.; Schmutz J.; Myers R.M.; Butterfield Y.S.;
RA Krzywinski M.I.; Skalska U.; Smalls D.E.; Schnerch A.; Schein J.E.;
RA Jones S.J.; Marra M.A.;
RP "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC059251; AAH59251.1; -.
DR InterPro; IPR000488; Death.
DR InterPro; IPR011029; DEATH_like.
DR InterPro; IPR000906; ZUS.
DR Pfam; PF00531; Death; 1.
DR Pfam; PF00791; ZUS; 1.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZUS; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
SQ SEQUENCE 1050 AA; 117435 MW; B0DD8C0591A8F177 CRC64;

Query Match 86.3%; Score 120; DB 2; Length 1050;
Best Local Similarity 92.0%; Pred. NO. 2.4e-08;
Matches 23; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ERIEERLAYIADHLGFSWTELARAL 26
|||||
Db 611 ERMEERLAYIADHLGFSWTELAREL 635

RESULT 6
Q8CCV0 PRELIMINARY; PRT; 1050 AA.
AC Q8CCV0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus adult male medulla oblongata cDNA, RIKEN full-length
DE enriched library, clone:633056N20 product:ankryrin 2, neuronal long
DE splice form homolog.
GN Name=Ank2;
```


[illegible]


```

Best Local Similarity 68.0%; Pred. NO. 0.00054;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ERIERELAYIADHGLGSWTSELARAL 26
   || : ||| ||| ||| ||| |||
Db 710 ERTDIRMAIVADHGLGSWTSELAR 734

Search completed: November 10, 2004, 12:27:26
Job time : 30.4475 secs

```

```

RESULT 14
Q6Z773 PRELIMINARY; PRT; 1101 AA.
ID Q6Z773
AC Q6Z773;
DT 05-JUN-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ44903.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Amysgdala;
RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
RA Arita M., Muesashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yanamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Masuo K., Nakamura Y., Sakine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Naganahri K., Masuho Y., Nagai K., Isogai T.,
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK126851; BAC86721.1; -.
DR InterPro; IPR000488; Death.
DR InterPro; IPR011029; DEATH_like.
DR InterPro; IPR00906; ZU5.
DR Pfam; PF00531; Death; 1.
DR Pfam; PF00791; ZU5; 1.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50017; DEATH_DOWAIN; 1.
SQ SEQUENCE 1101 AA; 12218 MW; 321372D4C0A60856 CRC64;
Query Match 64.7%; Score 90; DB 2; Length 1101;
Best Local Similarity 68.0%; Pred.No. 0.00054;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

```

QY      2  ERIEERLAYIADHGLGFSWTELAPAL 26
      || : || : || || || || || || ||
Db      710  ERTDIRMAIVADHGLGFSWTELAREL 734

RESULT 15
BAC86721
ID BAC86721 PRELIMINARY; PRT; 1101 AA.
AC BAC86721;
DT 02-MAR-2004 (TREMUREL. 27, Created)
DT 02-MAR-2004 (TREMUREL. 27, last sequence update)
DT 02-MAR-2004 (TREMUREL. 27, last annotation update)
DE CDNA FLJ44903 fis, clone BRAMY3005184, highly similar to Mus musculus
DE ankyrin 3, epithelial (Ank3).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [ ]
SEQUENCE FROM N.A.
RC TISSUE=Amnygdala;
RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
RA Arita M., Mutsaers K., Yuuki H., Hara H., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishi S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.
RA Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahara K., Masuho T., Nagai T.,
RT "NEDO human cDNA sequencing project.";
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
RW ENBL; AK126851; BAC86721.1;
SQ SEQUENCE 1101 AA; 122188 MW; 321372D4C0A60856 CRC64;

```

Query Match 64.7%; Score 90; DB 2; Length 1101;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:55:47 ; Search time 5.08997 Seconds
(without alignments)
143.349 Million cell updates/sec

Title: US-10-092-750-224
Perfect score: 63
Sequence: 1 KYQOLFEDIRW 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/prodata1/1/aa/5A COMB.pep.*
2: /cgn2_6/prodata1/1/aa/5B COMB.pep.*
3: /cgn2_6/prodata1/1/aa/6A COMB.pep.*
4: /cgn2_6/prodata1/1/aa/6B COMB.pep.*
5: /cgn2_6/prodata1/1/aa/PCTUS COMB.pep.*
6: /cgn2_6/prodata1/1/aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|------------------------|-------------------|
| 1 | 42 | 66.7 | 415 | 4 US-09-134-000C-3595 | Sequence 3595, Ap |
| 2 | 40 | 63.5 | 1980 | 4 US-09-914-272A-3 | Sequence 3, Appl |
| 3 | 39 | 61.9 | 365 | 4 US-10-039-659A-10 | Sequence 10, Appl |
| 4 | 39 | 61.9 | 374 | 3 US-09-045-583-48 | Sequence 48, Appl |
| 5 | 39 | 61.9 | 374 | 4 US-09-534-185-48 | Sequence 48, Appl |
| 6 | 39 | 61.9 | 549 | 3 US-09-291-922-30 | Sequence 30, Appl |
| 7 | 38 | 60.3 | 232 | 4 US-09-134-000C-4038 | Sequence 4038, Ap |
| 8 | 38 | 60.3 | 236 | 4 US-09-583-110-2857 | Sequence 2957, Ap |
| 9 | 38 | 60.3 | 347 | 3 US-08-960-780-25 | Sequence 25, Appl |
| 10 | 38 | 60.3 | 347 | 3 US-09-073-898-25 | Sequence 25, Appl |
| 11 | 38 | 60.3 | 347 | 4 US-09-307-106-6 | Sequence 6, Appl |
| 12 | 38 | 60.3 | 347 | 4 US-09-850-351A-25 | Sequence 25, Appl |
| 13 | 38 | 60.3 | 348 | 3 US-08-960-780-21 | Sequence 21, Appl |
| 14 | 38 | 60.3 | 348 | 3 US-08-960-780-42 | Sequence 42, Appl |
| 15 | 38 | 60.3 | 348 | 3 US-09-073-898-21 | Sequence 21, Appl |
| 16 | 38 | 60.3 | 348 | 3 US-09-073-898-42 | Sequence 42, Appl |
| 17 | 38 | 60.3 | 348 | 4 US-09-850-351A-21 | Sequence 21, Appl |
| 18 | 38 | 60.3 | 348 | 4 US-09-850-351A-42 | Sequence 42, Appl |
| 19 | 38 | 60.3 | 1290 | 1 US-08-138-641-2 | Sequence 2, Appl |
| 20 | 38 | 60.3 | 1290 | 1 US-08-138-133-2 | Sequence 2, Appl |
| 21 | 38 | 60.3 | 1290 | 4 US-09-538-092-956 | Sequence 956, App |
| 22 | 37 | 58.7 | 345 | 3 US-09-073-898-140 | Sequence 140, App |
| 23 | 37 | 58.7 | 345 | 4 US-09-307-106-34 | Sequence 34, Appl |
| 24 | 37 | 58.7 | 345 | 4 US-09-850-351A-140 | Sequence 140, App |
| 25 | 36 | 57.1 | 54 | 4 US-09-270-767-37516 | Sequence 37516, A |
| 26 | 36 | 57.1 | 54 | 4 US-09-270-767-52733 | Sequence 52733, A |
| 27 | 36 | 57.1 | 207 | 4 US-09-489-039A-13743 | Sequence 13743, A |

ALIGNMENTS

RESULT 1
US-09-134-000C-3595
; Sequence 3595, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134.000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3595
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3595

Query Match 66.7%; Score 42; DB 4; Length 415;
Best Local Similarity 80.0%; Pred. No. 21;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 KYQOLFEDIR 10
Db 131 KFOALFEDIR 140

RESULT 2
US-09-914-272A-3
; Sequence 3, Application US/09914272A
; Patent No. 6673913
; GENERAL INFORMATION:
; APPLICANT: Sakaguchi, No. 6673913uo
; APPLICANT: Kuwahara, Kazuhiko
; TITLE OF INVENTION: GANP Protein
; FILE REFERENCE: 050208-0014
; CURRENT APPLICATION NUMBER: US/09/914.272A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: PCT/JP99/04634
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: 47035/1999
; PRIOR FILING DATE: 1999-02-24
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1980
; TYPE: PRT

Oy 1 KYQOLFEDIR 10
Db 131 KFOALFEDIR 140

```
; ORGANISM: Homo sapiens
US-09-914-272A-3

Query Match      63.5%; Score 40; DB 4; Length 1980;
Best Local Similarity 60.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 YQQLFEDIRW 11
Db 1329 YQQLSDVAW 1338

RESULT 3
US-10-039-659A-10
; Sequence 10, Application US/10039659A
; Patent No. 6723520
; GENERAL INFORMATION:
; APPLICANT: Wang, Wei
; APPLICANT: Gish, Kurt C.
; APPLICANT: Schall, Thomas J.
; APPLICANT: Vicari, Alain P.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Antibodies that bind chemokine TECK
; FILE REFERENCE: DX0589KIB US
; CURRENT APPLICATION NUMBER: US/10/039,659A
; CURRENT FILING DATE: 2002-01-03
; PRIOR APPLICATION NUMBER: US 08/887,977
; PRIOR FILING DATE: 1997-07-03
; PRIOR APPLICATION NUMBER: US 60/021,664
; PRIOR FILING DATE: 1996-07-05
; PRIOR APPLICATION NUMBER: US 60/028,329
; PRIOR FILING DATE: 1996-10-11
; PRIOR APPLICATION NUMBER: US 60/048,593
; PRIOR FILING DATE: 1997-06-04
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-039-659A-10

Query Match      61.9%; Score 39; DB 4; Length 365;
Best Local Similarity 63.6%; Pred. No. 60;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KYQQLFEDIRW 11
Db 191 KYQTVSEPIRW 201

RESULT 4
US-09-045-583-48
; Sequence 48, Application US/09045583
; Patent No. 6287605
; GENERAL INFORMATION:
; APPLICANT: Graham, Gerard J. et al.
; TITLE OF INVENTION: No. 6287805el Molecules of the G Protein-Coupled
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/534,185
; FILING DATE: 24-Mar-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/045,583
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: MNI-044
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 374 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; APPLICATION NUMBER: US/09/045,583

; FILING DATE: 20-MAR-98
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: MNI-044
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 374 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal

Query Match      61.9%; Score 39; DB 3; Length 374;
Best Local Similarity 63.6%; Pred. No. 61;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KYQQLFEDIRW 11
Db 200 KYQTVSEPIRW 210

RESULT 5
US-09-534-185-48
; Sequence 48, Application US/09534185
; Patent No. 6403767
; GENERAL INFORMATION:
; APPLICANT: Graham, Gerard J. et al.
; TITLE OF INVENTION: No. 6403767el Molecules of the G Protein-Coupled
; Heptahelical Receptor Superfamily and Uses
; Therefor
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/534,185
; FILING DATE: 24-Mar-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/045,583
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: MNI-044
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 374 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
```

```
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-09-534-185-48
Query Match          61.9%; Score 39; DB 4; Length 374;
Best Local Similarity 63.6%; Pred. No. 61;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 KYQOLFEDIRW 11
Db      200 KYQTVSEPIRW 210

RESULT 6
US-09-291-922-30
; Sequence 30, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 30
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Beta vulgaris
US-09-291-922-30

Query Match          61.9%; Score 39; DB 3; Length 549;
Best Local Similarity 75.0%; Pred. No. 89;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      4 QLFEDIRW 11
Db      498 ELFEDFRW 505

RESULT 7
US-09-134-000C-4038
; Sequence 4038, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4038
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4038

Query Match          60.3%; Score 38; DB 4; Length 232;
Best Local Similarity 66.7%; Pred. No. 57;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 KYQOLFEDI 9
Db      43 QFQOLFEDV 51

;
; SEQUENCE DESCRIPTION: SEQ ID NO: 49:
US-09-583-110-2957
Query Match          60.3%; Score 38; DB 4; Length 236;
Best Local Similarity 70.0%; Pred. No. 58;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 KYQOLFEDIR 10
Db      3 KYQOLFQIQ 12

RESULT 9
US-08-960-780-25
; Sequence 25, Application US/08960780
; Patent No. 6204435
; GENERAL INFORMATION:
; APPLICANT: Feitelson, Jerald S.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stockhoff, Brian A.
; APPLICANT: Schmeits, James
; APPLICANT: Loewer, David
; APPLICANT: Dullum, Charles Joseph
; APPLICANT: Muller-Cohn, Judy
; APPLICANT: Stamp, Lisa
; TITLE OF INVENTION: No. 6204435a1 Pesticidal Toxins and Nucleotide
; TITLE OF INVENTION: Sequences Which Encode These Toxins
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/960,780
; APPLICATION NUMBER: US/08/960,780
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,848
; FILING DATE: 30-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
```

REFERENCE/DOCKET NUMBER: MA-708
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 347 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: 66D3
US-08-960-780-25

Query Match 60.3%; Score 38; DB 3; Length 347;
Best Local Similarity 54.5%; Pred. No. 84;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KYQQLFEDIRW 11
Db 38 KQOQYQSIW 48

RESULT 10

US-09-073-898-25
Sequence 25, Application US/09073898
Patent No. 6242669
GENERAL INFORMATION:
APPLICANT: Feitelson, Jerald S.
APPLICANT: Schnepf, H. Ernest
APPLICANT: Narva, Kenneth E.
APPLICANT: Stockhoff, Brian A.
APPLICANT: Schmeits, James
APPLICANT: Loewer, David
APPLICANT: Dullum, Charles Joseph
APPLICANT: Muller-Cohn, Judy
APPLICANT: Stamp, Lisa
APPLICANT: Morrill, George
APPLICANT: Finstad-Lee, Stacey
TITLE OF INVENTION: No. 6242669el Pesticidal Toxins and Nucleotide
TITLE OF INVENTION: Sequences Which Encode These Toxins
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606-6669
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,898
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/029,848
FILING DATE: 30-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/960,780
FILING DATE: 30-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-708C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:
LENGTH: 347 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: 66D3
US-09-073-898-25

Query Match 60.3%; Score 38; DB 3; Length 347;
Best Local Similarity 54.5%; Pred. No. 84;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KYQQLFEDIRW 11
Db 38 KQOQYQSIW 48

RESULT 11

US-09-307-106-6
Sequence 6, Application US/09307106
Patent No. 6603063
GENERAL INFORMATION:
APPLICANT: Feitelson, Jerald S.
APPLICANT: Schnepf, H. Ernest
APPLICANT: Narva, Kenneth E.
APPLICANT: Stockhoff, Brian A.
APPLICANT: Schmeits, James
APPLICANT: Loewer, David
APPLICANT: Dullum, Charles Joseph
APPLICANT: Muller-Cohn, Judy
APPLICANT: Stamp, Lisa
APPLICANT: Morrill, George
APPLICANT: Finstad-Lee, Stacey
TITLE OF INVENTION: No. 6603063el Pesticidal Toxins and Nucleotide
TITLE OF INVENTION: Sequences Which Encode These Toxins
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606-6669
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/307,106
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/029,848
FILING DATE: 30-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/960,780
FILING DATE: 30-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/073,898
FILING DATE: 05-MAY-1998
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-708C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 347 amino acids

```
/
/
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ ORIGINAL SOURCE: 66D3
/ INDIVIDUAL ISOLATE: 66D3
/ US-09-307-106-6

Query Match      60.3%; Score 38; DB 4; Length 347;
Best Local Similarity 54.5%; Pred. No. 84;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

INDIVIDUAL ISOLATE: 66D3
SEQUENCE DESCRIPTION: SEQ ID NO: 25;
US-09-850-351A-25

Query Match      60.3%; Score 38; DB 4; Length 347;
Best Local Similarity 54.5%; Pred. No. 84;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY      1 KYQQLFEDIRW 11
Db      38 KQOQTYQSIRW 48

RESULT 13
US-08-960-780-21
; Sequence 21, Application US/08960780
; Patent No. 6204435
; GENERAL INFORMATION:
; APPLICANT: Feitelson, Jerald S.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stockhoff, Brian A.
; APPLICANT: Schmeits, James
; APPLICANT: Loewer, David
; APPLICANT: Dullum, Charles Joseph
; APPLICANT: Muller-Cohn, Judy
; APPLICANT: Stamp, Lisa
; TITLE OF INVENTION: No. 6204435el Pesticidal Toxins and Nucleotide
; TITLE OF INVENTION: Sequences Which Encode These Toxins
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/960,780
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,848
; FILING DATE: 30-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA-708
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: 31J2
; US-08-960-780-21

Query Match      60.3%; Score 38; DB 3; Length 348;
Best Local Similarity 54.5%; Pred. No. 84;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY      1 KYQQLFEDIRW 11
Db      38 KQOQTYQSIRW 48

RESULT 12
US-09-850-351A-25
; Sequence 25, Application US/09850351A
; Patent No. 6656908
; GENERAL INFORMATION:
; APPLICANT: Feitelson, Jerald S.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stockhoff, Brian A.
; APPLICANT: Schmeits, James
; APPLICANT: Loewer, David
; APPLICANT: Dullum, Charles Joseph
; APPLICANT: Muller-Cohn, Judy
; APPLICANT: Stamp, Lisa
; APPLICANT: Morrill, George
; TITLE OF INVENTION: No. 6656908el Pesticidal Toxins and Nucleotide
; TITLE OF INVENTION: Sequences Which Encode These Toxins
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/850,351A
; FILING DATE: 07-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/073,898
; FILING DATE: 06-MAY-1998
; APPLICATION NUMBER: US 08/960,780
; FILING DATE: 30-OCT-1997
; APPLICATION NUMBER: US 60/029,848
; FILING DATE: 30-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
; REFERENCE/DOCKET NUMBER: MA-708CD1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
```

Db 39 KQOQTYQSIRW 49

RESULT 14

US-08-960-780-42
; Sequence 42, Application US/08960780
; Patent No. 6204435

GENERAL INFORMATION:

APPLICANT: Feitelson, Jerald S.
APPLICANT: Schnepf, H. Ernest
APPLICANT: Narva, Kenneth E.
APPLICANT: Stockhoff, Brian A.
APPLICANT: Schweits, James
APPLICANT: Loewer, David
APPLICANT: Dullum, Charles Joseph
APPLICANT: Muller-Cohn, Judy
APPLICANT: Stamp, Lisa

TITLE OF INVENTION: No. 6204435el Pesticidal Toxins and Nucleotide
NUMBER OF SEQUENCES: 134

CORRESPONDENCE ADDRESS:

ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/960,780
FILING DATE: 30-OCT-1997
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/029,848
FILING DATE: 30-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA-708

TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800

SEQUENCE CHARACTERISTICS:

LENGTH: 348 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: 197T1
US-08-960-780-42

Query Match 60.3%; Score 38; DB 3; Length 348;
Best Local Similarity 54.5%; Pred. No. 84;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KYOQLFEDIRW 11

Db 39 KQOQTYQSIRW 49

RESULT 15

US-09-073-898-21

; Sequence 21, Application US/09073898
; Patent No. 6242669

GENERAL INFORMATION:

APPLICANT: Feitelson, Jerald S.
APPLICANT: Schnepf, H. Ernest

APPLICANT: Narva, Kenneth E.
APPLICANT: Stockhoff, Brian A.
APPLICANT: Schweits, James
APPLICANT: Loewer, David
APPLICANT: Dullum, Charles Joseph
APPLICANT: Muller-Cohn, Judy
APPLICANT: Stamp, Lisa
APPLICANT: Morrill, George
APPLICANT: Finstad-Lee, Stacey
TITLE OF INVENTION: No. 6242669el Pesticidal Toxins and Nucleotide
NUMBER OF SEQUENCES: 144

CORRESPONDENCE ADDRESS:

ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,898
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/029,848
FILING DATE: 30-OCT-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/960,780
FILING DATE: 30-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-708C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800

SEQUENCE CHARACTERISTICS:

LENGTH: 348 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: 31J2
US-09-073-898-21

Query Match 60.3%; Score 38; DB 3; Length 348;
Best Local Similarity 54.5%; Pred. No. 84;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KYOQLFEDIRW 11

Db 39 KQOQTYQSIRW 49

Search completed: November 10, 2004, 15:57:23
Job time : 6.08897 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 02:43:24 ; Search time 26.5801 Seconds
(without alignments)
146.426 Million cell updates/sec

Title: US-10-092-750-224

Perfect score: 63

Sequence: 1 KYQLFEDIRW 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1568699 seqs, 353819137 residues

Total number of hits satisfying chosen parameters: 1568699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:*
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
 - 6: /cgn2_6/ptodata/2/pubpaa/ECTUS_PUBCOMB.pep:*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
 - 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
 - 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
 - 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
 - 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|----------------------|
| 1 | 63 | 100.0 | 11 | 14 | US-10-092-750-224 |
| 2 | 42 | 66.7 | 417 | 14 | US-10-369-493-3906 |
| 3 | 40 | 63.5 | 1086 | 15 | US-10-282-122A-74458 |
| 4 | 39 | 61.9 | 15 | 14 | US-10-239-423-45 |
| 5 | 39 | 61.9 | 30 | 14 | US-10-239-423-44 |
| 6 | 39 | 61.9 | 365 | 14 | US-10-039-659-10 |
| 7 | 39 | 61.9 | 365 | 16 | US-10-754-071-10 |
| 8 | 39 | 61.9 | 269 | 14 | US-10-239-423-69 |
| 9 | 39 | 61.9 | 374 | 14 | US-10-225-567A-390 |
| 10 | 39 | 61.9 | 374 | 14 | US-10-164-649-48 |
| 11 | 39 | 61.9 | 374 | 14 | US-10-239-423-68 |
| 12 | 39 | 61.9 | 390 | 13 | US-10-087-192-1920 |
| 13 | 39 | 61.9 | 406 | 14 | US-10-325-430-18 |

| | | | | | | |
|----|----|------|------|----|----------------------|-------------------|
| 14 | 39 | 61.9 | 406 | 14 | US-10-029-386-32911 | Sequence 32911, A |
| 15 | 39 | 61.9 | 548 | 10 | US-09-774-381-40 | Sequence 40, Appl |
| 16 | 39 | 61.9 | 549 | 13 | US-10-051-902-30 | Sequence 30, Appl |
| 17 | 39 | 61.9 | 549 | 13 | US-10-051-909-30 | Sequence 30, Appl |
| 18 | 38 | 60.3 | 147 | 17 | US-10-425-115-345726 | Sequence 345726, |
| 19 | 38 | 60.3 | 221 | 15 | US-10-282-122A-51899 | Sequence 51899, A |
| 20 | 38 | 60.3 | 236 | 9 | US-09-815-242-13450 | Sequence 13450, A |
| 21 | 38 | 60.3 | 236 | 15 | US-10-282-122A-74155 | Sequence 74155, A |
| 22 | 38 | 60.3 | 347 | 9 | US-09-850-351A-25 | Sequence 25, Appl |
| 23 | 38 | 60.3 | 347 | 14 | US-10-452-002A-6 | Sequence 6, Appl |
| 24 | 38 | 60.3 | 348 | 9 | US-09-850-351A-21 | Sequence 21, Appl |
| 25 | 38 | 60.3 | 348 | 9 | US-09-850-351A-42 | Sequence 42, Appl |
| 26 | 38 | 60.3 | 431 | 16 | US-10-437-963-177035 | Sequence 177035, |
| 27 | 38 | 60.3 | 477 | 14 | US-10-029-386-33072 | Sequence 33072, A |
| 28 | 38 | 60.3 | 504 | 14 | US-10-156-761-10228 | Sequence 10228, A |
| 29 | 38 | 60.3 | 1290 | 14 | US-10-017-128-2 | Sequence 2, Appl |
| 30 | 38 | 60.3 | 1290 | 16 | US-10-408-765A-477 | Sequence 477, App |
| 31 | 37 | 58.7 | 70 | 17 | US-10-425-115-360757 | Sequence 360757, |
| 32 | 37 | 58.7 | 74 | 15 | US-10-424-599-257432 | Sequence 257432, |
| 33 | 37 | 58.7 | 102 | 17 | US-10-425-115-342323 | Sequence 342323, |
| 34 | 37 | 58.7 | 127 | 15 | US-10-282-122A-65718 | Sequence 65718, A |
| 35 | 37 | 58.7 | 185 | 15 | US-10-108-260A-4160 | Sequence 4160, Ap |
| 36 | 37 | 58.7 | 204 | 15 | US-10-264-049-2729 | Sequence 2729, Ap |
| 37 | 37 | 58.7 | 307 | 9 | US-09-911-826A-7 | Sequence 7, Appl |
| 38 | 37 | 58.7 | 345 | 9 | US-09-850-351A-140 | Sequence 140, App |
| 39 | 37 | 58.7 | 345 | 14 | US-10-452-002A-34 | Sequence 34, Appl |
| 40 | 37 | 58.7 | 501 | 15 | US-10-424-593-176275 | Sequence 176275, |
| 41 | 37 | 58.7 | 599 | 15 | US-10-389-647-688 | Sequence 688, App |
| 42 | 36 | 57.1 | 77 | 14 | US-10-195-142-14 | Sequence 14, Appl |
| 43 | 36 | 57.1 | 121 | 16 | US-10-381-530-18 | Sequence 18, Appl |
| 44 | 36 | 57.1 | 126 | 15 | US-10-425-114-43219 | Sequence 43219, A |
| 45 | 36 | 57.1 | 129 | 16 | US-10-437-963-181535 | Sequence 181535, |

ALIGNMENTS

RESULT 1
US-10-092-750-224
; Sequence 224, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 224
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-224

Query Match 100.0%; Score 63; DB 14; Length 11;
Best Local Similarity 100.0%; Pred. NO. 0.00033;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KYQLFEDIRW 11
| | | | | | | | | | | |
Db 1 KYQLFEDIRW 11
| | | | | | | | | | | |

RESULT 2
US-10-369-493-3906
; Sequence 3906, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 3906
 ; LENGTH: 417
 ; TYPE: PRT
 ; ORGANISM: Neurospora crassa
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(417)
 ; OTHER INFORMATION: unsure at all Xaa locations
 US-10-369-493-3906

Query Match 66.7%; Score 42; DB 14; Length 417;
 Best Local Similarity 86.7%; Pred. No. 56;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 QQLFEDIRW 11
 :||:|
 Db 103 QRLFDLRW 111

RESULT 3
 US-10-282-122A-74458
 ; Sequence 74458, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 74458
 ; LENGTH: 1086
 ; TYPE: PRT
 ; ORGANISM: Streptococcus pyogenes
 US-10-282-122A-74458

Query Match 63.5%; Score 40; DB 15; Length 1086;
 Best Local Similarity 54.5%; Pred. No. 3.2e+02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KYQLFEDIRW 11
 |||:|
 Db 1063 KYQLFREFSW 1073

RESULT 4
 US-10-239-423-45
 ; Sequence 45, Application US/10239423
 ; Publication No. US20030186889A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FORSMANN, Wolf-Georg; FORSMANN, Ulf; ADERMAN, Knut;
 ; APPLICANT: HEITLAND, Aleksandra; SPODSBERG, Nikolaj
 ; TITLE OF INVENTION: Diagnostic Agent and Medicament for Examining the
 ; TITLE OF INVENTION: Cell Surface Proteome of Tumor and Inflammation Cells and
 ; TITLE OF INVENTION: for Treating Tumor Diseases and Inflammatory Diseases,
 ; TITLE OF INVENTION: Preferably with the Aid of Specific Chemokine
 ; TITLE OF INVENTION: Receptor Analysis and Chemokine Receptor/Ligand Interaction
 ; FILE REFERENCE: 022217us
 ; CURRENT APPLICATION NUMBER: US/10/239,423
 ; CURRENT FILING DATE: 2002-09-23
 ; PRIOR APPLICATION NUMBER: DE10016013.1
 ; PRIOR FILING DATE: 2000-03-31
 ; NUMBER OF SEQ ID NOS: 84
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 45
 ; LENGTH: 15
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:
 ; OTHER INFORMATION: Amino Acid Sequence for the Generation of Antibodies
 US-10-239-423-45

Query Match 61.8%; Score 39; DB 14; Length 15;
 Best Local Similarity 63.6%; Pred. No. 6.9;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KYQLFEDIRW 11
 |||:|
 Db 3 KYQTVSEPIW 13

RESULT 5
 US-10-239-423-44
 ; Sequence 44, Application US/10239423
 ; Publication No. US20030186889A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FORSMANN, Wolf-Georg; FORSMANN, Ulf; ADERMAN, Knut;
 ; APPLICANT: HEITLAND, Aleksandra; SPODSBERG, Nikolaj
 ; TITLE OF INVENTION: Diagnostic Agent and Medicament for Examining the
 ; TITLE OF INVENTION: Cell Surface Proteome of Tumor and Inflammation Cells and
 ; TITLE OF INVENTION: for Treating Tumor Diseases and Inflammatory Diseases,
 ; TITLE OF INVENTION: Preferably with the Aid of Specific Chemokine
 ; TITLE OF INVENTION: Receptor Analysis and Chemokine Receptor/Ligand Interaction
 ; FILE REFERENCE: 022217us
 ; CURRENT APPLICATION NUMBER: US/10/239,423
 ; CURRENT FILING DATE: 2002-09-23
 ; PRIOR APPLICATION NUMBER: DE10016013.1
 ; PRIOR FILING DATE: 2000-03-31
 ; NUMBER OF SEQ ID NOS: 84


```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: Amino Acid Sequence for the Generation of Antibodies
US-10-239-423-44

Query Match      61.9%; Score 39; DB 14; Length 30;
Best Local Similarity 63.6%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 KYQQLFEDIRW 11
      ||| : |||
Db      20 KYQTVSEPIRW 30

RESULT 6
US-10-039-659-10
; Sequence 10, Application US/10039659
; Publication No. US20030018167A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Wei
; APPLICANT: Gish, Kurt C.
; APPLICANT: Schall, Thomas J.
; APPLICANT: Vicari, Alain P.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: MAMMALIAN CHEMOKINE REAGENTS
; FILE REFERENCE: DX0589K1C US
; CURRENT APPLICATION NUMBER: US/10/039,659
; CURRENT FILING DATE: 2004-01-07
; PRIOR APPLICATION NUMBER: US 10/039,659
; PRIOR FILING DATE: 2002-01-03
; PRIOR APPLICATION NUMBER: US 08/887,977
; PRIOR FILING DATE: 1997-07-03
; PRIOR APPLICATION NUMBER: US 60/021,664
; PRIOR FILING DATE: 1996-07-05
; PRIOR APPLICATION NUMBER: US 60/028,329
; PRIOR FILING DATE: 1996-10-11
; PRIOR APPLICATION NUMBER: US 60/048,593
; PRIOR FILING DATE: 1997-06-04
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-754-071-10

Query Match      61.9%; Score 39; DB 16; Length 365;
Best Local Similarity 63.6%; Pred. No. 1.6e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 KYQQLFEDIRW 11
      ||| : |||
Db      191 KYQTVSEPIRW 201

RESULT 7
US-10-754-071-10
; Sequence 10, Application US/10754071
; Publication No. US20040137578A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Wei
; APPLICANT: Gish, Kurt C.
; APPLICANT: Schall, Thomas J.
; APPLICANT: Vicari, Alain P.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Chemokine TECK Polypeptides
; FILE REFERENCE: DX0589K1C US
; CURRENT APPLICATION NUMBER: US/10/754,071
; CURRENT FILING DATE: 2004-01-07
; PRIOR APPLICATION NUMBER: US 10/039,659
; PRIOR FILING DATE: 2002-01-03
; PRIOR APPLICATION NUMBER: US 08/887,977
; PRIOR FILING DATE: 1997-07-03
; PRIOR APPLICATION NUMBER: US 60/021,664
; PRIOR FILING DATE: 1996-07-05
; PRIOR APPLICATION NUMBER: US 60/028,329
; PRIOR FILING DATE: 1996-10-11
; PRIOR APPLICATION NUMBER: US 60/048,593
; PRIOR FILING DATE: 1997-06-04
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-754-071-10

Query Match      61.9%; Score 39; DB 16; Length 365;
Best Local Similarity 63.6%; Pred. No. 1.6e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 KYQQLFEDIRW 11
      ||| : |||
Db      191 KYQTVSEPIRW 201

RESULT 8
US-10-239-423-69
; Sequence 69, Application US/10239423
; Publication No. US20030186889A1
; GENERAL INFORMATION:
; APPLICANT: FORSSMANN, Wolf-Georg; FORSSMANN, Ulf; ADERMAN, Knut;
; APPLICANT: HEITLAND, Aleksandra; SPODSBERG, Nikolaj
; TITLE OF INVENTION: Diagnostic Agent and Medicament for Examining the
; TITLE OF INVENTION: Cell Surface Proteome of Tumor and Inflammation Cells and
; TITLE OF INVENTION: for Treating Tumor Diseases and Inflammatory Diseases,
; TITLE OF INVENTION: Preferably with the Aid of Specific Chemokine
; TITLE OF INVENTION: Receptor Analysis and Chemokine Receptor/Ligand Interaction
; FILE REFERENCE: 022217us
; CURRENT APPLICATION NUMBER: US/10/239,423
; CURRENT FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: DE10016013.1
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 69
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: Amino Acid Sequence for the Generation of Antibodies
US-10-039-659-10

Query Match      61.9%; Score 39; DB 14; Length 365;
Best Local Similarity 63.6%; Pred. No. 1.6e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 KYQQLFEDIRW 11
      ||| : |||
Db      20 KYQTVSEPIRW 30

RESULT 6
US-10-039-659-10
; Sequence 10, Application US/10039659
; Publication No. US20030018167A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Wei
; APPLICANT: Gish, Kurt C.
; APPLICANT: Schall, Thomas J.
; APPLICANT: Vicari, Alain P.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: MAMMALIAN CHEMOKINE REAGENTS
; FILE REFERENCE: DX0589K1C US
; CURRENT APPLICATION NUMBER: US/10/039,659
; CURRENT FILING DATE: 2004-01-07
; PRIOR APPLICATION NUMBER: US 10/039,659
; PRIOR FILING DATE: 2002-01-03
; PRIOR APPLICATION NUMBER: US 08/887,977
; PRIOR FILING DATE: 1997-07-03
; PRIOR APPLICATION NUMBER: US 60/021,664
; PRIOR FILING DATE: 1996-07-05
; PRIOR APPLICATION NUMBER: US 60/028,329
; PRIOR FILING DATE: 1996-10-11
; PRIOR APPLICATION NUMBER: US 60/048,593
; PRIOR FILING DATE: 1997-06-04
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-754-071-10

Query Match      61.9%; Score 39; DB 16; Length 365;
Best Local Similarity 63.6%; Pred. No. 1.6e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 KYQQLFEDIRW 11
      ||| : |||
Db      191 KYQTVSEPIRW 201

RESULT 7
US-10-754-071-10
; Sequence 10, Application US/10754071
; Publication No. US20040137578A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Wei
; APPLICANT: Gish, Kurt C.
; APPLICANT: Schall, Thomas J.
; APPLICANT: Vicari, Alain P.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Chemokine TECK Polypeptides
; FILE REFERENCE: DX0589K1C US
; CURRENT APPLICATION NUMBER: US/10/754,071
; CURRENT FILING DATE: 2004-01-07
; PRIOR APPLICATION NUMBER: US 10/039,659
; PRIOR FILING DATE: 2002-01-03
; PRIOR APPLICATION NUMBER: US 08/887,977
; PRIOR FILING DATE: 1997-07-03
; PRIOR APPLICATION NUMBER: US 60/021,664
; PRIOR FILING DATE: 1996-07-05
; PRIOR APPLICATION NUMBER: US 60/028,329
; PRIOR FILING DATE: 1996-10-11
; PRIOR APPLICATION NUMBER: US 60/048,593
; PRIOR FILING DATE: 1997-06-04
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-754-071-10

Query Match      61.9%; Score 39; DB 16; Length 365;
Best Local Similarity 63.6%; Pred. No. 1.6e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 KYQQLFEDIRW 11
      ||| : |||
Db      191 KYQTVSEPIRW 201

RESULT 8
US-10-239-423-69
; Sequence 69, Application US/10239423
; Publication No. US20030186889A1
; GENERAL INFORMATION:
; APPLICANT: FORSSMANN, Wolf-Georg; FORSSMANN, Ulf; ADERMAN, Knut;
; APPLICANT: HEITLAND, Aleksandra; SPODSBERG, Nikolaj
; TITLE OF INVENTION: Diagnostic Agent and Medicament for Examining the
; TITLE OF INVENTION: Cell Surface Proteome of Tumor and Inflammation Cells and
; TITLE OF INVENTION: for Treating Tumor Diseases and Inflammatory Diseases,
; TITLE OF INVENTION: Preferably with the Aid of Specific Chemokine
; TITLE OF INVENTION: Receptor Analysis and Chemokine Receptor/Ligand Interaction
; FILE REFERENCE: 022217us
; CURRENT APPLICATION NUMBER: US/10/239,423
; CURRENT FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: DE10016013.1
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 69
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: Amino Acid Sequence for the Generation of Antibodies
US-10-039-659-10
```

```

US-10-239-423-69
;
; REFERENCE/DOCKET NUMBER: MMI-044
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 374 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-10-164-649-48
Query Match 61.9%; Score 39; DB 14; Length 374;
Best Local Similarity 63.6%; Pred. No. 1.7e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KYQOLFEDIRW 11
Db 195 KYQTVSEPIRW 205
|||||:|||||

RESULT 9
US-10-225-567A-390
; Sequence 390, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 390
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-567A-390

Query Match 61.9%; Score 39; DB 14; Length 374;
Best Local Similarity 63.6%; Pred. No. 1.7e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KYQOLFEDIRW 11
Db 200 KYQTVSEPIRW 210
|||||:|||||

RESULT 10
US-10-164-649-48
; Sequence 48, Application US/10164649
; Publication No. US20030162943A1
; GENERAL INFORMATION:
; APPLICANT: Graham, Gerard J. et al.
; TITLE OF INVENTION: No. US20030162943A1el Molecules of the G Protein-Coupled
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/164,649
; FILING DATE: 07-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/045,583
; FILING DATE: 20-MAR-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragoras, Amy E.
; REGISTRATION NUMBER: 36,207

```

; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1920
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-087-192-1920

Query Match 61.9%; Score 39; DB 13; Length 390;
Best Local Similarity 63.6%; Pred. No. 1.7e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KYQOLFEDIRW 11
|||:||||
Db 216 KYQTVSEPIRW 226

RESULT 13
US-10-325-430-18
; Sequence 18, Application US/10325430
; Publication No. US20030153525A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: Rosenfeld, Julie Beth
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 1465, 1587, 2146, 2207,
; TITLE OF INVENTION: 32838,336 and 52908
; FILE REFERENCE: MPI01-294PIRW
; CURRENT APPLICATION NUMBER: US/10/325,430
; PRIOR FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US 60/341,953
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-325-430-18

Query Match 61.9%; Score 39; DB 14; Length 406;
Best Local Similarity 63.8%; Pred. No. 1.8e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KYQOLFEDIRW 11
|||:||||
Db 232 KYQTVSEPIRW 242

RESULT 14
US-10-029-386-32911
; Sequence 32911, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 32911
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

; OTHER INFORMATION: MAP TO AL121935.12
; OTHER INFORMATION: EXPRESSED IN LONG, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.8
; OTHER INFORMATION: SWISSPROT HIT: P51684, EVALUATE 0.00e+00
US-10-029-386-32911

Query Match 61.9%; Score 39; DB 14; Length 406;
Best Local Similarity 63.6%; Pred. No. 1.8e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KYQOLFEDIRW 11
|||:||||
Db 232 KYQTVSEPIRW 242

RESULT 15
US-09-774-381-40
; Sequence 40, Application US/09774381
; Publication No. US20030082677A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Pan, Yang
; APPLICANT: Gearing, David P.
; TITLE OF INVENTION: MTR-1, LSP-1, TAP-1, AND PA-I MOLECULES
; TITLE OF INVENTION: AND US9S THEREFOR
; FILE REFERENCE: MWI-107CP2
; CURRENT APPLICATION NUMBER: US/09/774,381
; CURRENT FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 08/941,354
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 09/010,674
; PRIOR FILING DATE: 1998-01-22
; PRIOR APPLICATION NUMBER: 60/061,149
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 09/014,347
; PRIOR FILING DATE: 1998-01-27
; PRIOR APPLICATION NUMBER: 60/061,159
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 09/474,151
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 09/004,206
; PRIOR FILING DATE: 1998-01-08
; PRIOR APPLICATION NUMBER: 60/061,143
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 09/483,414
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 09/213,571
; PRIOR FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 08/994,890
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Beta vulgaris
US-09-774-381-40

Query Match 61.9%; Score 39; DB 10; Length 548;
Best Local Similarity 75.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 QLFEDIRW 11
:|||||
Db 497 ELFEDFRW 504

Search completed: November 11, 2004, 07:41:46
Job time : 26:5801 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:52:32 ; Search time 3.75801 Seconds
(without alignments)
281.634 Million cell updates/sec

Title: US-10-092-750-224

Perfect score: 63

Sequence: 1 KYQLFEDIRW 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 52 | 82.5 | 862 | 1 S56766 | replication licens |
| 2 | 52 | 82.5 | 863 | 1 S65954 | replication licens |
| 3 | 50 | 79.4 | 858 | 2 T47223 | replication licens |
| 4 | 50 | 79.4 | 863 | 1 S64720 | replication licens |
| 5 | 42 | 66.7 | 266 | 2 C95316 | probable ABC trans |
| 6 | 40 | 63.5 | 1872 | 2 T00339 | hypothetical prote |
| 7 | 39 | 61.9 | 162 | 2 JCS068 | conserved hypothet |
| 8 | 39 | 61.9 | 369 | 2 D69899 | G protein-coupled |
| 9 | 39 | 61.9 | 549 | 2 T14606 | probable sugar tra |
| 10 | 38 | 60.3 | 159 | 2 S14122 | hypothetical prote |
| 11 | 38 | 60.3 | 221 | 2 D97161 | probable enzyme wi |
| 12 | 38 | 60.3 | 236 | 2 D95220 | trehalose operon t |
| 13 | 38 | 60.3 | 236 | 2 B98084 | hypothetical prote |
| 14 | 38 | 60.3 | 310 | 2 T41158 | hypothetical prote |
| 15 | 38 | 60.3 | 629 | 2 T14776 | 1-phosphatidylinos |
| 16 | 38 | 60.3 | 1290 | 2 A36466 | 1-phosphatidylinos |
| 17 | 38 | 60.3 | 1230 | 2 A31317 | 1-phosphatidylinos |
| 18 | 38 | 60.3 | 1231 | 2 S00666 | conserved hypothet |
| 19 | 37 | 58.7 | 127 | 2 H81021 | conserved hypothet |
| 20 | 37 | 58.7 | 381 | 2 E69352 | hypothetical prote |
| 21 | 37 | 58.7 | 589 | 2 A42385 | hypothetical prote |
| 22 | 37 | 58.7 | 599 | 2 E83241 | hypothetical prote |
| 23 | 36 | 57.1 | 108 | 1 MNIV2F | nonstructural prot |
| 24 | 36 | 57.1 | 118 | 1 MNIVX2 | nonstructural prot |
| 25 | 36 | 57.1 | 118 | 1 MNIVX4 | nonstructural prot |
| 26 | 36 | 57.1 | 118 | 1 MNIVX6 | nonstructural prot |
| 27 | 36 | 57.1 | 118 | 1 MNIVX8 | nonstructural prot |
| 28 | 36 | 57.1 | 118 | 1 MNIVXX | nonstructural prot |
| 29 | 36 | 57.1 | 121 | 1 E45539 | nonstructural prot |

30 57.1 121 1 MNIV2 nonstructural prot
31 57.1 121 1 MNIV2A nonstructural prot
32 57.1 121 1 MNIV2M nonstructural prot
33 57.1 121 1 MNIV2W nonstructural prot
34 57.1 121 1 MNIV62 nonstructural prot
35 57.1 121 1 MNIVB1 nonstructural prot
36 57.1 121 1 MNIVB3 nonstructural prot
37 57.1 121 1 MNIVB4 nonstructural prot
38 57.1 121 1 MNIVB5 nonstructural prot
39 57.1 121 1 MNIVB7 nonstructural prot
40 57.1 121 1 MNIVB8 nonstructural prot
41 57.1 121 1 MNIVC2 nonstructural prot
42 57.1 121 1 MNIV26 nonstructural prot
43 57.1 121 1 MNIV2K nonstructural prot
44 57.1 121 1 MNIVB6 nonstructural prot
45 57.1 121 2 S09649 NS2 protein - infl

ALIGNMENTS

RESULT 1

S56766

replication licensing factor MCM4 - mouse

N/Alternate names: cell division cycle control protein CDC21/CDC54

C/Species: Mus musculus (house mouse)

C/Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004

C/Accession: S56766

R/Kimura, H.; Takizawa, N.; Nozaki, N.; Sugimoto, K.

Nucleic Acids Res. 23, 2097-2104, 1995

A/Title: Molecular cloning of cDNA encoding mouse Cdc21 and CDC46 homologs and characteri

A/Reference number: S56766; MUID:95334361; PMID:7610039

A/Accession: S56766

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: mRNA

A/Residues: 1-862 <KM>

A/Cross-references: UNIPROT:P49717; EMBL:D26089; NID:9940405; PIDN:BA05082.1; PID:G94040

A/Note: The nucleotide sequence was submitted to the EMBL Data Library, December 1993

C/Comment: The complex of six MCM proteins is one of several proteins that must be bound

phosphorylated and dissociate from the chromatin.

C/Complex: The predominant form is a heterohexamer of MCM2, MCM3, MCM4, MCM5, MCM6, and M

C/Function:

A/Description: part of the replication licensing system that permits DNA replication to c

C/Superfamily: replication licensing factor MCM4; MCM homology

C/Keywords: cell cycle control; DNA replication initiation; heterohexamer; nucleus; phos

F:270-763/Domain: MCM homology <MCM>

Query Match Similarity 82.5%; Score 52; DB 1; Length 862;

Best Local Similarity 100.0%; Pred. No. 0.33; Mismatches 0; Indels 0; Gaps 0;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KYQLFEDIR 10

|||||

Db 818 KYQLFEDIR 827

RESULT 2

S65954

replication licensing factor MCM4 - human

N/Alternate names: cell division cycle control protein CDC21/CDC54; S. cerevisiae minich

C/Species: Homo sapiens (man)

C/Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004

C/Accession: S65954; S43198; S41622

R/Musahl, C.; Schulte, D.; Burkhardt, R.; Knippers, R.

Eur. J. Biochem. 230, 1096-1101, 1995

A/Title: A human homologue of the yeast replication protein Cdc21. Interactions with othe

A/Reference number: S65954; MUID:95324568; PMID:7601140

A/Accession: S65954

A/Status: nucleic acid sequence not shown

A/Molecule type: mRNA

A/Residues: 'IHERVATSSASGRIPRNARRGRALARRRPRCGAQGSRRAGPARACPCRAGTST',1-863 <MUS>

A/Cross-references: UNIPROT:P33991; EMBL:X74794; NID:9683749; PIDN:CAA52801.1; PID:G9405

R/Hu, B.

submitted to the EMBL Data Library, August 1993

A:Reference number: S43198
A:Accession: S43198
A:Molecule type: mRNA
A:Residues: 'ARE', 440-702 <HUB1>
A:Cross-references: EMBL:X74794
R:Hu, B.; Burkhardt, R.; Schulte, D.; Musahl, C.; Knippers, R.
Nucleic Acids Res. 21, 5289-5293, 1993
A:Title: The P1 family: a new class of nuclear mammalian proteins related to the yeast M
A:Reference number: S41622; MUID:94089373; PMID:8265339
A:Accession: S41622
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 'ARE', 440-636 <HUB2>
A:Cross-references: EMBL:X74794
A:Comment: The complex of six MCM proteins is one of several proteins that must be bound
phosphorylated and dissociate from the chromatin.
C:Genetics:
A:Gene: GDB:MCM4; CDC21; CDC54
A:Cross-references: GDB:433798
C:Complex: The predominant form is a heterohexamer of MCM2, MCM3, MCM4, MCM5, MCM6, and
C:Function:
A:Description: part of the replication licensing system that permits DNA replication to
C:Superfamily: replication licensing factor MCM4; MCM homology
C:Keywords: cell cycle control; DNA replication initiation; heterohexamer; nucleus; phos
F:271-764/Domain: MCM homology <MCM>

Query Match 82.5%; Score 52; DB 1; Length 863;

Best Local Similarity 100.0%; Pred. No. 0.34; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KYQQLFEDIR 10

Db 819 KYQQLFEDIR 828

RESULT 3

T47223
A:Description: replication licensing factor MCM4 (validated) - African clawed frog
N:Alternate names: cdc21p
C:Species: Xenopus laevis (African clawed frog)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C:Accession: T47223
R:Kubota, Y.; Mimura, S.; Nishimoto, S.; Masuda, T.; Nojima, H.; Takisawa, H.
EMBO J. 16, 3320-3331, 1997
A:Title: Licensing of DNA replication by a multi-protein complex of MCM/p1 proteins in X
A:Reference number: 224400; MUID:97357318; PMID:9214647
A:Accession: T47223
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-858 <KUB>
A:Cross-references: UNIPROT:O42589; EMBL:U44049; NID:g2231172; PIDN:AAAC60225.1; PID:g223
C:Complex: the replication licensing complex consists of at least six proteins; MCM2 (PI
alicated, MUID:97357318)
C:Function:
A:Description: MCM4 is a component of the replication licensing factor that permits DNA
A:Superfamily: replication licensing factor MCM4; MCM homology
F:266-759/Domain: MCM homology <MCM>

Query Match 79.4%; Score 50; DB 2; Length 858;

Best Local Similarity 90.0%; Pred. No. 0.76;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KYQQLFEDIR 10

Db 814 KYQQLFEDIR 823

RESULT 4

S64720
A:Description: replication licensing factor MCM4 - African clawed frog
N:Alternate names: cell division cycle control protein CDC21/CDC54
C:Species: Xenopus laevis (African clawed frog)

C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: S64720; S26643; S25529
R:Coue, M.; Kearsley, S.E.; Mechali, M.
EMBO J. 15, 1085-1097, 1996
A:Title: Chromatin binding, nuclear localization and phosphorylation of Xenopus cdc21 are
A:Reference number: S64720; MUID:96183193; PMID:8605878

A:Accession: S64720

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-863 <COU>

A:Cross-references: UNIPROT:P30664; EMBL:U29178; NID:gl002597; PIDN:AA801680.1; PID:gl002
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995
R:Coxon, A.; Maundrell, K.; Kearsley, S.E.
Nucleic Acids Res. 20, 5571-5577, 1992

A:Title: Fission yeast cdc21(+) belongs to a family of proteins involved in an early step

A:Reference number: S26640; MUID:93087163; PMID:1454522

A:Accession: S26643

A:Molecule type: DNA

A:Residues: 513-523, 'Y', 525-538, 'G', 540-553, 'G', 555-588 <COX>

A:Cross-references: EMBL:Z15033; NID:964612; PIDN:CAA78751.1; PID:964613

C:Comment: The complex of six MCM proteins is one of several proteins that must be bound
phosphorylated and dissociate from the chromatin.

C:Genetics:

A:Gene: cdc21

A:Complex: The predominant form is a heterohexamer of MCM2, MCM3, MCM4, MCM5, MCM6, and
C:Function:

A:Description: part of the replication licensing system that permits DNA replication to
C:Superfamily: replication licensing factor MCM4; MCM homology

C:Keywords: cell cycle control; DNA replication initiation; heterohexamer; nucleus; phos
F:271-764/Domain: MCM homology <MCM>

Query Match 79.4%; Score 50; DB 1; Length 863;

Best Local Similarity 90.0%; Pred. No. 0.77;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KYQQLFEDIR 10

Db 819 KYQQLFEDIR 828

RESULT 5

C95316
A:Description: probable ABC transporter, permease protein Sma0802 [imported] - Sinorhizobium meliloti
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: C95316
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.;
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti

A:Reference number: A95262; MUID:21396509; PMID:11481432

A:Accession: C95316

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-266 <KUR>

A:Cross-references: UNIPROT:Q92ZM0; GB:AB006469; PIDN:AAK65093.1; PID:gl4523529; GSPDB:GN

A:Experimental source: strain 1021, megaplasmid pSymA

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: Sma0802

A:Genome: plasmid

C:Superfamily: spermidine/putrescine transport system permease protein potI

Query Match 66.7%; Score 42; DB 2; Length 266;

Best Local Similarity 70.0%; Pred. No. 5.7;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 YQQLFEDIRW 11
||:|||||
Db 56 YQELLEDFRW 65

RESULT 6
T00339
hypothetical protein KIAA0572 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C:Accession: T00339
R:Kagase, T.; Isinikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, O.
DNA Res. 5, 31-39, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. IX. The completed
A:Reference number: Z14086; MUID:98290545; PMID:9628581
A:Accession: T00339
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1872 <NAG>
A:Cross-references: UNIPROT:O60318; EMBL:AB011144; NID:G3043667; PIDN:BAA25498.1; PID:G3
A:Experimental source: brain; clone HH2391
C:Genetics:
A:Note: KIAA0572

Query Match 63.5%; Score 40; DB 2; Length 1872;
Best Local Similarity 60.0%; Pred. No. 1; Indels 0;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 YQQLFEDIRW 11
|||:|:
Db 1221 YQQLSDVAW 1230

RESULT 7
D69899
conserved hypothetical protein yobM - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: D69899
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
A.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
C.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrati, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Kotter, P.; Konigstein, C.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardiniois,
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: D69899
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-162 <KUN>
A:Cross-references: UNIPROT:O34377; GB:Z99114; GB:AL009126; NID:G2634230; PIDN:CAB13793.
A:Experimental source: strain 168
C:Genetics:
A:Gene: yobM

Query Match 61.9%; Score 39; DB 2; Length 162;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KYQQLFEDI 9
||:|||||
Db 79 KYEDLFEDI 87

RESULT 8
JC5068
G Protein-coupled receptor CKR-L3 - human
C:Species: Homo sapiens (man)
C>Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 21-Jul-2000
C:Accession: JC5068
R:Zaballos, A.; Varona, R.; Gutierrez, J.; Lind, P.; Marquez, G.
Biochem. Biophys. Res. Commun. 227, 846-853, 1996
A:Title: Molecular cloning and RNA expression of two new human chemokine receptor-like g
A:Reference number: JC5067; MUID:97040707; PMID:8886020
A:Accession: JC5068
A:Molecule type: DNA
A:Residues: 1-369 <ZAB>
A:Cross-references: EMBL:Z79784; NID:g1668737; PIDN:CA802144.1; PID:g1668738
C:Comment: This protein belongs to the family of alpha chemokine receptors.
C:Genetics:
A:Gene: GDB:CMKBR6; STRL22; GPR29; CCR6; CKR-L3; GPR-CY4
A:Cross-references: GDB:15370639; OMIM:601835
A:Map position: 6q27-6q27
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; transmembrane protein
F:42-68/Domain: transmembrane #status predicted <TM1>
F:79-99/Domain: transmembrane #status predicted <TM2>
F:115-136/Domain: transmembrane #status predicted <TM3>
F:160-180/Domain: transmembrane #status predicted <TM4>
F:212-233/Domain: transmembrane #status predicted <TM5>
F:250-271/Domain: transmembrane #status predicted <TM6>
F:292-315/Domain: transmembrane #status predicted <TM7>

Query Match 61.9%; Score 39; DB 2; Length 369;
Best Local Similarity 63.6%; Pred. No. 28;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KYQQLFEDIRW 11
|||:|:
Db 195 KYQTVSEPIRW 205

RESULT 9
TI4606
probable sugar transport protein 205 - beet
C:Species: Beta vulgaris (beet)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: TI4606; TI4617
R:Chlou, T.J.; Bush, D.R.
Plant Physiol. 110, 511-520, 1996
A:Title: Molecular cloning, immunochemical localization to the vacuole, and expression in
A:Reference number: Z18131; MUID:96351183; PMID:8742332
A:Accession: TI4606
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-549 <CHI>
A:Cross-references: UNIPROT:P93075; EMBL:U64902; NID:g1778092; PID:g1778093
A:Accession: TI4617
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-541, 'SVQV' <CH2>
A:Cross-references: EMBL:U64903; NID:g1778094; PID:g1778095
C:Genetics:
A:Note: Bvcdna-205; Bvcdna-397
C:Superfamily: glucose transport protein
C:Keywords: sugar transport

Query Match 61.9%; Score 39; DB 2; Length 549;
Best Local Similarity 75.0%; Pred. No. 44;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 QLFEDIRW 11
||:|||||
Db 498 ELFEDEFW 505

RESULT 10

SI4122
 hypothetical protein B (COXIII 5' region) - common sunflower mitochondrion
 C:Species: mitochondrion Helianthus annuus (common sunflower)
 C>Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 09-Jul-2004
 C:Accession: SI4122; S52011; S48855
 R:Quagliariello, C.; Salardi, A.; Gallerani, R.
 Curr. Genet. 18, 355-363, 1990
 A:Title: The cytochrome oxidase subunit III gene in sunflower mitochondria is cotranscribed
 A:Reference number: SI4122; MUID:91070622; PMID:2174745
 A:Accession: SI4122
 A:Molecule type: DNA
 A:Residues: 1-159 <QUA>
 A:Cross-references: UNIPROT:P41248; EMBL:X57669; NID:G450343; PIDN:CAA40864.1; PID:G4438
 R:Spasova, M.; Moninger, F.; Leaver, C.J.; Petrov, P.; Atanasov, A.; Nijkamp, H.J.J.; H
 Plant Mol. Biol. 26, 1819-1831, 1994
 A:Title: Characterisation and expression of the mitochondrial genome of a new type of cy
 A:Reference number: S52010; MUID:95161706; PMID:7858220
 A:Accession: S52011
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-159 <SPA>
 A:Cross-references: EMBL:X82386; NID:G563544; PIDN:CAA57787.1; PID:G563546
 C:Genetics:
 A:Genome: mitochondrion
 C:Superfamily: wheat mitochondrial 18K membrane protein.
 C:Keywords: mitochondrion
 F:1-85/Domain: H+-transporting ATP synthase alpha chain homology (fragment) <ATP>

Query Match 60.3%; Score 38; DB 2; Length 159;
 Best Local Similarity 60.0%; Pred. No. 17;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 YQQLFEDIRW 11
 |||||:
 Db 79 YSLFSDSQW 88

RESULT 11

D97161
 probable enzyme with a TIM-barrel fold [imported] - Clostridium acetobutylicum
 C:Species: Clostridium acetobutylicum
 C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
 C:Accession: D97161
 R:Noll, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
 J. Bacteriol. 183, 4823-4838, 2001
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl
 A:Reference number: A96900; MUID:21359325; PMID:21359325
 A:Accession: D97161
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-221 <KUR>
 A:Cross-references: UNIPROT:Q97H92; GB:AE001437; PIDN:AAK80079.1; PID:G15025111; GSPDB:G
 A:Experimental source: Clostridium acetobutylicum ATCC824
 C:Genetics:
 A:Gene: CAC2121
 C:Superfamily: conserved hypothetical protein HI0090

Query Match 60.3%; Score 38; DB 2; Length 221;
 Best Local Similarity 45.5%; Pred. No. 24;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 KYQQLFEDIRW 11
 |||||:
 Db 54 KYDEMNSDVRW 64

RESULT 12

D95220
 trehalose operon transcription repressor [imported] - Streptococcus pneumoniae (strain T
 C:Species: Streptococcus pneumoniae

C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
 C:Accession: D95220
 R:Testelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heide
 on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, F
 son, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
 A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; MUID:21357209; PMID:11463916

A:Accession: D95220
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-236 <KUR>
 A:Cross-references: UNIPROT:Q97NW8; UNIPROT:Q8DN15; GB:AE005672; PIDN:AAK75957.1; PID:G14
 A:Experimental source: strain TIGR4
 C:Genetics:
 A:Gene: SPI885

Query Match 60.3%; Score 38; DB 2; Length 236;
 Best Local Similarity 70.0%; Pred. No. 26;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KYQQLFEDIR 10
 |||||:
 Db 3 KYQQLFKQIQ 12

RESULT 13

B98084
 trehalose operon transcription repressor [imported] - Streptococcus pneumoniae (strain R6
 C:Species: Streptococcus pneumoniae
 C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
 C:Accession: B98084
 R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
 e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
 Y, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
 A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A:Reference number: A97872; MUID:21429245; PMID:11544234
 A:Accession: B98084
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-236 <KUR>
 A:Cross-references: UNIPROT:Q97NW8; UNIPROT:Q8DN15; GB:AE007317; PIDN:AAI00503.1; PID:G14
 C:Genetics:
 A:Gene: trer

Query Match 60.3%; Score 38; DB 2; Length 236;
 Best Local Similarity 70.0%; Pred. No. 26;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KYQQLFEDIR 10
 |||||:
 Db 3 KYQQLFKQIQ 12

RESULT 14

T41158
 hypothetical protein SPCCL8.15 - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: T41158
 R:Hilbert, H.; Duesterhoeft, A.; Wood, V.; Rajadream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, October 1998

A:Reference number: Z21973
 A:Accession: T41158
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-310 <HIL>
 A:Cross-references: UNIPROT:O74865; EMBL:AL031907; PIDN:CAA21429.1; GSPDB:GN00068; SPDB:S
 A:Experimental source: strain 972h-; cosmid c18
 C:Genetics:

A;Gene: SPDB:SPCC18.15
A;Map position: 3
A;Introns: 59/2

Query Match 60.3%; Score 38; DB 2; Length 310;
Best Local Similarity 66.7%; Pred. No. 35;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KYQQLFEDI 9
DB 22 KYSQVFEDV 30

RESULT 15

Tl4776
Hypothetical protein DKFZp434N101.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: Tl4776
R;Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, August 1999
A;Reference number: Z18183
A;Accession: Tl4776
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-629 <POU>
A;Cross-references: UNIPROT:Q3UFY1; EMBL:AL110247
A;Experimental source: adult testis; clone DKFZp434N101
C;Genetics:
A;Note: DKFZp434N101.1
C;Superfamily: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase II; 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase II; SH3 homology

Query Match 60.3%; Score 38; DB 2; Length 629;
Best Local Similarity 70.0%; Pred. No. 77;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KYQQLFEDIR 10
DB 538 RYQQPFEDFR 547

Search completed: November 10, 2004, 15:55:08
Job time : 4.75801 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:50:40 ; Search time 18.4377 Seconds
(without alignments)
343.270 Million cell updates/sec

Title: US-10-092-750-224
Perfect score: 63
Sequence: 1 KYQQLFEDIRW 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|---------------------|
| 1 | 52 | 82.5 | 862 | 1 MCM4_MOUSE | P49717 mus musculus |
| 2 | 52 | 82.5 | 862 | 2 Q8C120 | Q8C120 mus musculus |
| 3 | 52 | 82.5 | 862 | 2 Q9D077 | Q9D077 mus musculus |
| 4 | 52 | 82.5 | 862 | 2 Q921D5 | Q921D5 mus musculus |
| 5 | 52 | 82.5 | 862 | 2 BAC40578 | Bac40578 mus muscu |
| 6 | 52 | 82.5 | 863 | 1 MCM4_HUMAN | P33991 homo sapien |
| 7 | 52 | 82.5 | 863 | 2 Q8NEH1 | Q8NEH1 homo sapien |
| 8 | 52 | 82.5 | 863 | 2 AAS83108 | Aas83108 homo sapi |
| 9 | 50 | 79.4 | 858 | 2 Q42589 | Q42589 xenopus lae |
| 10 | 50 | 79.4 | 863 | 1 MCM4_XENLA | Q6GL41 xenopus tro |
| 11 | 50 | 79.4 | 863 | 2 Q6GL41 | Q6GL41 xenopus tro |
| 12 | 47 | 74.6 | 845 | 2 Q6NZV2 | Q6NZV2 brachydanio |
| 13 | 47 | 74.6 | 845 | 2 AAh65958 | Aah65958 brachydan |
| 14 | 45 | 71.4 | 911 | 2 Q975J8 | Q975J8 sulfolobus |
| 15 | 44 | 69.8 | 500 | 2 Q7MA20 | Q7MA20 wolinnella s |
| 16 | 42 | 66.7 | 266 | 2 Q922M0 | Q922M0 rhizobium m |
| 17 | 42 | 66.7 | 486 | 2 Q834N6 | Q834N6 enterococcu |
| 18 | 41 | 65.1 | 125 | 2 Q7QMZ1 | Q7QMZ1 anopheles g |
| 19 | 40 | 63.5 | 110 | 2 Q7N8Q5 | Q7N8Q5 photorhabd |
| 20 | 40 | 63.5 | 286 | 2 Q6EZ27 | Q6EZ27 campylobact |
| 21 | 40 | 63.5 | 302 | 2 Q9Z421 | Q9Z421 pseudomonas |
| 22 | 40 | 63.5 | 705 | 2 Q825D9 | Q825D9 nitrosomona |
| 23 | 40 | 63.5 | 778 | 2 Q6PJPE | Q6PJPE homo sapien |
| 24 | 40 | 63.5 | 778 | 2 AAh13285 | Aah13285 homo sapi |
| 25 | 40 | 63.5 | 1086 | 2 Q938K3 | Q938K3 temperate p |
| 26 | 40 | 63.5 | 1086 | 2 Q8K5Q1 | Q8K5Q1 streptococc |
| 27 | 40 | 63.5 | 1086 | 2 Q8NZS4 | Q8NZS4 streptococc |
| 28 | 40 | 63.5 | 1086 | 2 Q9A0N0 | Q9A0N0 streptococc |
| 29 | 40 | 63.5 | 1980 | 1 MCM4_HUMAN | O60318 homo sapien |
| 30 | 40 | 63.5 | 1980 | 2 AAS89300 | Aas89300 homo sapi |
| 31 | 39 | 61.9 | 162 | 2 Q34377 | Q34377 bacillus su |

ALIGNMENTS

RESULT 1

MCM4_MOUSE

ID MCM4_MOUSE STANDARD; PRT; 862 AA.

AC P49717; O89056;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE DNA replication licensing factor MCM4 (CDC21 homolog) (P1-CDC21).
GN Name=Mcm4; Synonyms=Mcm4, Cdc21;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
EX MEDLINE=95334361; PubMed=7610039;
RA Kimura H., Takizawa N., Nozaki N., Sugimoto K.;
RT "Molecular cloning of cDNA encoding mouse Cdc21 and CDC46 homologs and
RT characterization of the products: physical interaction between
RT P1(MCM3) and CDC46 proteins.";
RL Nucleic Acids Res. 23:2097-2104(1995).
RN [2]
RP SEQUENCE OF 503-602 FROM N.A.
RC STRAIN=BALB/c; TISSUE=Spleen;
RX MEDLINE=99012997; PubMed=9798653;
RA Chu C.C., Paul W.E.;
RT "Expressed genes in interleukin-4 treated B cells identified by cDNA
RT representational difference analysis.";
RL Mol. Immunol. 35:487-502(1998).
CC -!- FUNCTION: Involved in the control of DNA replication.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the MCM family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; D26089; BAA05082.1; -;
CC EMBL; U89402; AAC36509.1; -;
CC PIR; S56766; S56766.
CC MGD; MGI:103199; Mcm4.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR001208; MCM.
CC InterPro; IPR008047; MCM_4.
CC Pfam; PF00493; MCM; 1.
CC PRINTS; PR01657; MCMFAMILY.
CC PRINTS; PR01660; MCMPROTEIN4.
CC ProDom; PD001041; MCM; 1.
CC SMART; SM00382; AAA; 1.
CC SMART; SM00350; MCM; 1.

Q6MNM4 bdellovibri
Cae79127 bdellovib
P51684 homo sapien
Q6CPJ4 kluyveromyc
P93076 beta vulgar
P93075 beta vulgar
Q7RKF0 plasmodium
Q8ZTF7 pyrobaculum
P41248 heliantus
Q97H92 streptidum
Q8DN15 streptococc
Q97NW8 streptococc
Q6W249 rhizobium s
Q7VGH4 helicobacte

DR PROSITE; PS00847; MCM_1; 1.
 DR PROSITE; PSS0051; MCM_2; 1.
 KW ATP-binding; DNA replication; DNA-binding; Nuclear protein;
 KW Transcription regulation.
 FT DOMAIN 457 666 MCM.
 FT NP BIND 509 516 ATP (Potential).
 FT NP BIND 530 530 Q -> R (in Ref. 2).
 FT CONFLICT 530 530 I -> T (in Ref. 2).
 FT CONFLICT 572 572 I -> T (in Ref. 2).
 SQ SEQUENCE 862 AA; 96736 MW; 516ACCI3AC8FB16E CRC64;

Query Match 82.5%; Score 52; DB 1; Length 862;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KYQQLFEDIR 10
 |||||
 Db 818 KYQQLFEDIR 827

RESULT 2
 ID Q8C120 PRELIMINARY; PRT; 862 AA.
 AC Q8C120
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Mus musculus brain CRL-1443 BC3H1 CDNA, RIKEN full-length enriched
 DE library, clone:G430050N09 product:mini chromosome maintenance
 DE deficient 4 homolog (S. cerevisiae), full insert sequence.
 GN Name=Mcm4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H; TISSUE=Brain;
 RX MEDLINE=9279253; PubMed=10349636;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning."
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H; TISSUE=Brain;
 RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H; TISSUE=Brain;
 RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H; TISSUE=Brain;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P., Shibata K., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes."
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H; TISSUE=Brain;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,

RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer."
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H; TISSUE=Brain;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito K., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK089999; BAC41036.1; --
 DR MGD; MGI:103199; MCM4.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0008094; F:DNA-dependent ATPase activity; IEA.
 DR GO; GO:0006270; P:DNA replication initiation; IEA.
 DR InterPro; IP000345; CytC_heme_BS.
 DR InterPro; IP000208; MCM.
 DR InterPro; IP0008047; MCM_4.
 DR Pfam; PF00493; MCM; 1.
 DR PRINTS; PR01657; MCMFAMILY.
 DR PRINTS; PR01660; MCMPROTEIN4.
 DR ProDom; PD01041; MCM; 1.
 DR SMART; SM00350; MCM; 1.
 DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE; PS00051; MCM_2; 1.
 SQ SEQUENCE 862 AA; 96752 MW; FF2FA597021EBB56 CRC64;

Query Match 82.5%; Score 52; DB 2; Length 862;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KYQQLFEDIR 10
 |||||
 Db 818 KYQQLFEDIR 827

RESULT 3
 ID Q9D077 PRELIMINARY; PRT; 862 AA.
 AC Q9D077
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Mus musculus 10 days embryo whole body cDNA, RIKEN full-length
 DE enriched library, clone:2610042115 product:mini chromosome maintenance
 DE deficient 4 homolog (S. cerevisiae), full insert sequence.
 GN Name=Mcm4;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=9279253; PubMed=10349636;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning."
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX STRAIN=C57BL/6J; TISSUE=Whole body;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20493374; PubMed=11042159;
RA Carninci P, Shibata Y, Hayatsu N, Sugahara Y, Shibata K, Itoh M.,
RA Konno H, Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RL Prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX STRAIN=C57BL/6J; TISSUE=Whole body;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Ktsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA Adachi J., Aizawa K., Akhira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoaka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi P., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AK011743; BA027813.1; -.
DR MGD; MGI:103199; Mcm4.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008094; F:DNA-dependent ATPase activity; IEA.
DR GO; GO:000166; F:nucleotide binding; IEA.
DR GO; GO:0006270; P:DNA replication initiation; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001208; MCM.
DR InterPro; IPR008047; MCM_4.
DR Pfam; PF00493; MCM; 1.
DR PRINTS; PR01657; MCMFAMILY.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS00847; MCM_1; 1.
DR PROSITE; PS0051; MCM_2; 1.
DR ATP-binding.

SQ SEQUENCE 862 AA; 96764 MW; 6533BD3FA75A7866 CRC64;
Query Match 82.5%; Score 52; DB 2; Length 862;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KYOOLFEDIR 10
Db 818 KYOOLFEDIR 827
RESULT 4
Q921D5 PRELIMINARY; PRT; 862 AA.
AC Q921D5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Minichromosome maintenance protein 4.
GN Name=Mcm4;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N;
RC TISSUE=Mammary tumor. Metallothionien-TGF alpha model. 10 month old
RC virgin mouse. Taken by biopsy.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnertch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N;
RC TISSUE=Mammary tumor. Metallothionien-TGF alpha model. 10 month old
RC virgin mouse. Taken by biopsy.
RA Strausberg R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC013094; AAHL3094.1; -.
DR MGD; MGI:103199; Mcm4.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008094; F:DNA-dependent ATPase activity; IEA.
DR GO; GO:000166; F:nucleotide binding; IEA.
DR GO; GO:0006270; P:DNA replication initiation; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001208; MCM.
DR InterPro; IPR008047; MCM_4.
DR Pfam; PF00493; MCM; 1.
DR PRINTS; PR01657; MCMFAMILY.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS00847; MCM_1; 1.
DR PROSITE; PS0051; MCM_2; 1.
DR ATP-binding.

Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
"RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer.";
Genome Res. 10:1757-1771(2000).
[6]
SEQUENCE FROM N.A.
STRAIN=NOJ; TISSUE=Thymus;
Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
Hayashida K., Hayatsu N., Hiramoto K., Hiraoaka T., Hirozane T.,
Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
Kurihara C., Matsuyama T., Mizazaki A., Murata M., Nakamura M.,
Nishii K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
Saito K., Saitoh K., Sakai C., Sakai K., Sakazume N., Sano H.,
Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
EMBL; AK088796; BAC40578.1; -- 516ACCLAC3C6PB15E CFC64;
SEQUENCE 862 AA; 96735 MW; 516ACCLAC3C6PB15E CFC64;
Query Match 82.5%; Score 52; DB 2; Length 862;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KYQOLFEDIR 10
DB 818 KYQOLFEDIR 827

RESULT 6
MCM4_HUMAN
IID MCM4_HUMAN STANDARD; PRT; 863 AA.
P33991; Q99658;
01-FEB-1994 (Rel. 28, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
05-JUL-2004 (Rel. 44, Last annotation update)
DNA replication licensing factor MCM4 (CDC21 homolog) (P1-CDC21).
Name=MCM4; Synonyms=CDC21;
Homo sapiens (Human)
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1] _
SEQUENCE FROM N.A.
Hu B.;
Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE OF 1-712 FROM N.A.
MEDLINE=98126438; PubMed=9465298;
Connelly M.A., Zhang H., Kieleczawa J., Anderson C.W.;
"The promoters for human DNA-PKcs (PRKDC) and MCM4: divergently
transcribed genes located at chromosome 8 band q11.";
Genomics 47:71-83(1998).
[3]
SEQUENCE OF 1-23 FROM N.A.
MEDLINE=97430835; PubMed=9284934;
Ladenburger E.M., Packelmayer P.O., Hamsteier H., Knippers R.;
"MCM4 and PRKDC, human genes encoding proteins MCM4 and DNA-PKcs, are
close neighbours located on chromosome 8q12--q13.";
Cytogenet. Cell Genet. 77:268-270(1997).
[4]
SEQUENCE OF 440-863 FROM N.A.
TISSUE=Cervix;
MEDLINE=94089373; PubMed=8265339;
Hu B., Burkhardt R., Schulte D., Musahl C., Knippers R.;
"The p1 family: a new class of nuclear mammalian proteins related to
the yeast Mcm replication proteins.";
Nucleic Acids Res. 21:5289-5293(1993).
-!- FUNCTION: Involved in the control of DNA replication.
-!- SUBCELLULAR LOCATION: Nuclear (By similarity).
-!- SIMILARITY: Belongs to the MCM family.

RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY588245; AAC83108.1; -. 6D79DA81DD580BAF CRC64;
SQ SEQUENCE 863 AA; 96575 MW; 5D79DA81DD580BAF CRC64;

Query Match 82.5%; Score 52; DB 2; Length 863;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KYQQLFEDIR 10

Db 819 KYQQLFEDIR 828

RESULT 9

ID O42589 PRELIMINARY; PRT; 858 AA.

AC O42589;

DT 01-JAN-1998 (TrEMBLrel. 05, Created)

DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Cdc21p.

OS Xenopus laevis (African clawed frog).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

CC Xenopodinae; Xenopus.

OX NCBI_TaxID=8355;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97357318; PubMed=92114647;

RA Kubota Y., Mimura S., Nishimoto S., Masuda T., Nojima H., Takisawa H.;

RT "Licensing of DNA replication by a multi-protein complex of MCM/P1

proteins in Xenopus eggs."

RL EMBL J. 16:3320-3331(1997).

DR EMBL; U44049; AAC60225.1; -. PIR; T47223; T47223.

DR GO; GO:0005634; C:nucleus; IEA.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0003677; F:DNA binding; IEA.

DR GO; GO:0008994; F:DNA-dependent ATPase activity; IEA.

DR GO; GO:0000166; F:nucleotide binding; IEA.

DR GO; GO:0006270; P:DNA replication initiation; IEA.

DR InterPro; IPR003593; AAA_ATPase.

DR InterPro; IPR001208; MCM.

DR InterPro; IPR008047; MCM_4.

DR Pfam; PF00493; MCM_1.

DR PRINTS; PR01657; MCMFAMILY.

DR ProDom; PD01660; MCMFAMILY.

DR SMART; SM001041; MCM; 1.

DR SMART; SM00382; AAA; 1.

DR SMART; SM00350; MCM; 1.

DR PROSITE; PS00847; MCM_1; 1.

DR PROSITE; PS50051; MCM_2; 1.

KW ATP-binding.

SQ SEQUENCE 858 AA; 97053 MW; 0ACBFACT2943AF34 CRC64;

Query Match 79.4%; Score 50; DB 2; Length 858;

Best Local Similarity 90.0%; Pred. No. 3.4;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KYQQLFEDIR 10

Db 814 KYQQLFEDIR 823

RESULT 10

MCM4_XENLA

ID MCM4_XENLA STANDARD; PRT; 863 AA.

AC P30664; Q91679;

DT 01-APR-1993 (Rel. 25, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE DNA replication licensing factor MCM4 (CDC21 homolog) (P1-CDC21)

DE (X.MCM4).

GN Name=MCM4;
OS Xenopus laevis (African clawed frog).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
CC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96183193; PubMed=8605878;
RA Coue M., Kearsey S.E., Mechali M.;
RT "Chromatin binding, nuclear localization and phosphorylation of
RT Xenopus cdc21 are cell-cycle dependent and associated with the control
RT of initiation of DNA replication."
RL EMBL J. 15:1085-1097(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97057222; PubMed=8901561;
RA Hendrickson M., Madine M., Dalton S., Gautier J.;
RT "Phosphorylation of MCM4 by cdc2 protein kinase inhibits the activity
RT of the minichromosome maintenance complex.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:12223-12228(1996).
RN [3]
RP SEQUENCE OF 513-588 FROM N.A.
RX MEDLINE=93087163; PubMed=1454522;
RA Coxon A., Maundrell K., Kearsey S.E.;
RT "Fission yeast cdc21+ belongs to a family of proteins involved in an
RT early step of chromosome replication."
RL Nucleic Acids Res. 20:5571-5577(1992).
CC -!- FUNCTION: Involved in the control of DNA replication.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the MCM family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U29178; AAB01680.1; -. PIR; S64720; S64720.
DR EMBL; U46131; AAA91232.1; -. PIR; S64720; S64720.
DR EMBL; Z15033; CAA78751.1; -. PIR; S64720; S64720.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001208; MCM.
DR InterPro; IPR008047; MCM_4.
DR Pfam; PF00493; MCM; 1.
DR PRINTS; PR01657; MCMFAMILY.
DR PRINTS; PR01660; MCMFAMILY.
DR ProDom; PD001041; MCM; 1.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00350; MCM; 1.
DR PROSITE; PS00847; MCM_1; 1.
DR PROSITE; PS50051; MCM_2; 1.
KW ATP-binding; DNA replication; DNA-binding; Nuclear protein;
KW Transcription regulation.
FT DOMAIN 458 667 MCM.
FT NP BIND 510 517 ATP (Potential).
FT NP BIND 517 573 V -> A (in Ref. 2).
FT CONFLICT 73 73 RA -> LS (in Ref. 2).
FT CONFLICT 740 741
SQ SEQUENCE 863 AA; 97178 MW; 9A47003097EEDE2E CRC64;

Query Match 79.4%; Score 50; DB 1; Length 863;

Best Local Similarity 90.0%; Pred. No. 3.4;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KYQQLFEDIR 10

Db 819 KYQQLFEDIR 828

RESULT 11

Q6GL41
ID Q6GL41 PRELIMINARY; PRT; 863 AA.
AC Q6GL41
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8364;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC074670; AAH74670.1; -
DR InterPro; IPR001208; MCM.
DR Pfam; PF00493; MCM; 1.
DR PRINTS; PR01657; MCMFAMILY.
DR ProDom; PD001041; MCM; 1.
DR SMART; SM00350; MCM; 1.
DR PROSITE; PS00847; MCM_1; 1.
DR PROSITE; PS00851; MCM_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 863 AA; 97195 MW; 2AAFF32FE85BE97B CRC64;

Query Match 79.4%; Score 50; DB 2; Length 863;
Best Local Similarity 90.0%; Pred. No. 3.4;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KYQQLFEDIR 10
Db 819 KYQQLFEDIR 828
RESULT 12
Q6NZV2 PRELIMINARY; PRT; 845 AA.
ID Q6NZV2
AC Q6NZV2
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Mcm4 protein.
GN Name=mcm4;
OS Brachydanio rerio (Zebrafish) (Danio rerio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Strausberg R.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC065958; AAH65958.1; -
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001208; MCM.
DR InterPro; IPR008047; MCM_4.
DR Pfam; PF00493; MCM; 1.
DR PRINTS; PR01657; MCMFAMILY.
DR ProDom; PD001041; MCM; 1.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00350; MCM; 1.
DR PROSITE; PS00847; MCM_1; 1.
DR PROSITE; PS00851; MCM_2; 1.
KW ATP-binding.
SQ SEQUENCE 845 AA; 94523 MW; 8D959D2C956A63DA CRC64;

Query Match 74.6%; Score 47; DB 2; Length 845;
Best Local Similarity 80.0%; Pred. No. 12;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KYQQLFEDIR 10
Db 801 KYQQLFEDIR 810
RESULT 13
AAH65958 PRELIMINARY; PRT; 845 AA.
ID AAH65958
AC AAH65958
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;

RA MEDLINE=22388257; PubMed=12477932;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshlyki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Basa S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Strausberg R.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
ENBL: BC065958; AAH65958.1; -
KW Hypothetical protein.
SQ SEQUENCE 845 AA; 94523 MW; 8D959D2C956A63DA CRC64;

Query Match 74.6%; Score 47; DB 2; Length 845;
Best Local Similarity 80.0%; Pred. No. 12;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KYQOLFEDIR 10
DB 801 KYQOLFEDIR 810
|||||:|:
|||:|:

RESULT 14
QY75L8 PRELIMINARY; PRT; 911 AA.
ID QY75L8
AC QY75L8
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 911aa long hypothetical ATP-dependent helicase.
GN OrderedLocusNames=ST0401;
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=11955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX MEDLINE=21456156; PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankaei A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Okina T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermoacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7";
RL DNA Res. 8:123-140(2001).
RL ENBL: AP000982; BAB65382.1; -
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO: GO:0016787; F:hydrolyase activity; IEA.
DR GO: GO:0003676; F:nucleic acid binding; IEA.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR011545; DEAD/DEAH_N.
DR InterPro: IPR001650; Helicase_C.

DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Complete proteome; Helicase; Hydrolase;
KW Hypothetical protein.
SQ SEQUENCE 911 AA; 104448 MW; 24B758D9F066200C CRC64;

Query Match 71.4%; Score 45; DB 2; Length 911;
Best Local Similarity 54.5%; Pred. No. 30;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 KYQOLFEDIR 11
DB 135 KYQOLFEDIR 145
|||||:|:
|||:|:

RESULT 15
QYMA20 PRELIMINARY; PRT; 500 AA.
ID QYMA20
AC QYMA20;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ATP SYNTHASE F1 ALPHA SUBUNIT (EC 3.6.3.14).
GN Name=ATPA; OrderedLocusNames=WS0514;
OS Wolinella succinogenes.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Wolinella.
OX NCBI_TaxID=844;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSMZ 1740;
RX MEDLINE=22882897; PubMed=14500908;
RA Baar C., Epinger M., Radatz G., Simon J., Lanz C., Klimmek O.,
RA Nandakumar R., Gross R., Rosinus A., Keller H., Jagtap P., Linke B.,
RA Meyer F., Lederer H., Schuster S.C.;
RT "Complete genome sequence and analysis of Wolinella succinogenes";
RL Proc. Natl. Acad. Sci. U.S.A. 100:11690-11695(2003).
CC -!- FUNCTION: Produces ATP from ADP in the presence of an ion gradient
CC across the membrane (By similarity).
CC -!- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic
CC core - and CF(0) - the membrane proton channel. CF(1) has five
CC subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0)
CC has three main subunits: a, b and c (By similarity).
CC -!- SIMILARITY: Belongs to the ATPase alpha/beta chains family.
DR ENBL: BX571658; CAB09651.1; -
DR GO: GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.
DR GO: GO:0046961; F:hydrogen-transporting ATPase activity; IEA.
DR GO: GO:0016787; F:hydrolyase activity; IEA.
DR GO: GO:0016820; F:hydrolyase activity; IEA.
DR GO: GO:0015986; P:ATP synthesis coupled proton transport; IEA.
DR InterPro: IPR000194; ATPase_a/bcentre.
DR InterPro: IPR000793; ATPase_a/b_C.
DR InterPro: IPR004100; ATPase_a/b_N.
DR InterPro: IPR005294; ATP_synthF_alph.
DR Pfam; PF00006; ATP-synt_ab; 1.
DR Pfam; PF00306; ATP-synt_ab_C; 1.
DR Pfam; PF02874; ATP-synt_ab_N; 1.
DR TIGRFAMs; TIGR00962; atpA; 1.
DR PROSITE; PS00152; ATPASE ALPHA BETA; 1.
KW ATP synthetase; ATP-binding; CF(1); Complete proteome; Hydrolase;
KW Ion transport; Transport.
SQ SEQUENCE 500 AA; 54763 MW; F5F44084FFEF6D32 CRC64;

Query Match 69.8%; Score 44; DB 2; Length 500;
Best Local Similarity 80.0%; Pred. No. 24;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KYQOLFEDIR 10
|||||:|:
|||:|:

Db 465 KYPQIFEDIR 474

Search completed: November 10, 2004, 15:53:37
Job time : 20.5146 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:55:47 ; Search time 7.40214 Seconds
(without alignments)
143.349 Million cell updates/sec

Title: US-10-092-750-225

Perfect score: 76
Sequence: 1 IGEEFSRAAEKLYLA 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata1/iaa/5A COMB pep.*
- 2: /cgn2_6/prodata1/iaa/5B COMB pep.*
- 3: /cgn2_6/prodata1/iaa/6A COMB pep.*
- 4: /cgn2_6/prodata1/iaa/6B COMB pep.*
- 5: /cgn2_6/prodata1/iaa/PCUS COMB pep.*
- 6: /cgn2_6/prodata1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 43 | 56.6 | 155 | 4 | US-09-270-767-33093 |
| 2 | 43 | 56.6 | 155 | 4 | US-09-270-767-48310 |
| 3 | 43 | 56.6 | 175 | 2 | US-08-997-080-53 |
| 4 | 43 | 56.6 | 175 | 2 | US-08-997-362-53 |
| 5 | 43 | 56.6 | 175 | 3 | US-08-873-970-53 |
| 6 | 43 | 56.6 | 175 | 3 | US-09-095-855-53 |
| 7 | 43 | 56.6 | 175 | 3 | US-08-705-347A-53 |
| 8 | 43 | 56.6 | 175 | 3 | US-09-324-542-53 |
| 9 | 43 | 56.6 | 175 | 4 | US-09-205-426-53 |
| 10 | 43 | 56.6 | 175 | 4 | US-09-200-643-53 |
| 11 | 43 | 56.6 | 708 | 4 | US-09-602-787A-314 |
| 12 | 43 | 56.6 | 713 | 4 | US-09-602-787A-312 |
| 13 | 42 | 55.3 | 312 | 4 | US-09-252-991A-18082 |
| 14 | 42 | 55.3 | 324 | 4 | US-09-328-352-6612 |
| 15 | 42 | 55.3 | 536 | 4 | US-08-538-092-1309 |
| 16 | 41 | 53.9 | 283 | 4 | US-09-710-279-332 |
| 17 | 41 | 53.9 | 274 | 3 | US-09-134-001C-3921 |
| 18 | 40 | 52.6 | 149 | 4 | US-09-270-767-49034 |
| 19 | 40 | 52.6 | 298 | 4 | US-09-489-039A-8600 |
| 20 | 40 | 52.6 | 308 | 4 | US-09-543-681A-7814 |
| 21 | 40 | 52.6 | 984 | 4 | US-09-252-991A-17368 |
| 22 | 39 | 51.3 | 99 | 4 | US-08-562-991A-30614 |
| 23 | 39 | 51.3 | 101 | 2 | US-08-580-388A-21 |
| 24 | 39 | 51.3 | 106 | 2 | US-08-464-517-25 |
| 25 | 39 | 51.3 | 106 | 2 | US-08-246-361A-25 |
| 26 | 39 | 51.3 | 106 | 3 | US-08-463-772-25 |
| 27 | 39 | 51.3 | 106 | 5 | PCT-US93-05000-25 |

ALIGNMENTS

RESULT 1

US-09-270-767-33093

; Sequence 33093, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 33093

; LENGTH: 155

; TYPE: PRT

; ORGANISM: *Drosophila melanogaster*

US-09-270-767-33093

Query Match 56.6%; Score 43; DB 4; Length 155;

Best Local Similarity 53.3%; Pred. No. 4.5;

Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 IGEEFSRAAEKLYLA 15

:|||:|:|||||

Db 27 LGENFEQMEQLYLA 41

RESULT 2

US-09-270-767-48310

; Sequence 48310, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 48310

; LENGTH: 155

; TYPE: PRT

; ORGANISM: *Drosophila melanogaster*

US-09-270-767-48310

Query Match 56.6%; Score 43; DB 4; Length 155;

Best Local Similarity 53.3%; Pred. No. 4.5;

Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 IGEEFSRAAEKLYLA 15

Sequence 3, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 4, Appli
Sequence 4400, Ap
Sequence 30475, A
Sequence 18421, A
Sequence 8, Appli
Sequence 8, Appli
Sequence 965, App
Sequence 7212, Ap
Sequence 7, Appli
Sequence 7, Appli
Sequence 7, Appli
Sequence 31524, A
Sequence 8001, Ap

28 39 51.3 150 2 US-08-460-694-3
29 39 51.3 150 3 US-08-460-744-3
30 39 51.3 150 3 US-07-667-711B-3
31 39 51.3 173 1 US-08-193-977-4
32 39 51.3 282 4 US-09-543-681A-4400
33 39 51.3 344 4 US-09-252-991A-30475
34 39 51.3 411 4 US-09-252-991A-18421
35 39 51.3 432 1 US-08-522-166-8
36 39 51.3 432 1 US-08-488-382A-8
37 39 51.3 432 2 US-08-480-912-8
38 39 51.3 432 4 US-09-538-092-965
39 39 51.3 486 4 US-09-107-532A-7212
40 39 51.3 819 2 US-08-464-517-7
41 39 51.3 819 2 US-08-246-361A-7
42 39 51.3 819 3 US-08-463-772-7
43 39 51.3 819 5 PCT-US93-05000-7
44 38 50.0 243 4 US-09-252-991A-31524
45 38 50.0 284 4 US-09-489-039A-8001

Db 27 LGENFEQMEQLYLA 41
: || | : | : || |

```

RESULT 3
US-08-997-080-53
; Sequence 53, Application US/08997080
; Patent No. 5968524
; GENERAL INFORMATION:
; APPLICANT: WATSON, JAMES D.
; APPLICANT: TAN, PAUL L.J.
; TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,080
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 175 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-997-080-53

Query Match 56.6%; Score 43; DB 2; Length 175;
Best Local Similarity 75.0%; Pred. No. 5.1;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GEEFSRAAEKLY 13
| | | | |
Db 18 GYEFSPACEALY 29

RESULT 4
US-08-997-362-53
; Sequence 53, Application US/08997362
; Patent No. 5985287
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Hiwama, Jun
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Scott, Linda
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:

```

```

; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,362
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
; FILING DATE: June 12, 1997
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347
; FILING DATE: August 29, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002c2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 175 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-997-362-53

Query Match 56.6%; Score 43; DB 2; Length 175;
Best Local Similarity 75.0%; Pred. No. 5.1;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GEEFSRAAEKLY 13
| | | | |
Db 18 GYEFSPACEALY 29

RESULT 5
US-08-873-970-53
; Sequence 53, Application US/08873970
; Patent No. 6001361
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Hiwama, Jun
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Scott, Linda
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/873,970
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/705,347
; FILING DATE: 29-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 175 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-873-970-53
Query Match 56.6%; Score 43; DB 3; Length 175;
Best Local Similarity 75.0%; Pred. No. 5.1;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GEEFSRAAEKLY 13
| | | | |
Db 18 GYEFSSRAAEALY 29

RESULT 6
US-09-095-855-53
; Sequence 53, Application US/09095855
; Patent No. 6160093
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Hiyama, Jun
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Compounds and Methods for
; TITLE OF INVENTION: Treatment and Diagnosis of Mycobacterial Infections
; NUMBER OF SEQUENCES: 208
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,855
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/705,347
; FILING DATE: 29-AUG-1996
; APPLICATION NUMBER: 08/873,970
; FILING DATE: 12-JUN-1997
; APPLICATION NUMBER: 08/997,362
; FILING DATE: 23-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565

; APPLICATION NUMBER: US/08/873,970
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/705,347
; FILING DATE: 29-AUG-1996
; APPLICATION NUMBER: 08/873,970
; FILING DATE: 12-JUN-1997
; APPLICATION NUMBER: 08/997,362
; FILING DATE: 23-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 175 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-095-855-53
Query Match 56.6%; Score 43; DB 3; Length 175;
Best Local Similarity 75.0%; Pred. No. 5.1;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GEEFSRAAEKLY 13
| | | | |
Db 18 GYEFSSRAAEALY 29

RESULT 7
US-08-705-347A-53
; Sequence 53, Application US/08705347A
; Patent No. 6284255
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Hiyama, Jun
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Scott, Linda
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Speckman Picard PLLC
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/705,347A
; FILING DATE: 28-AUG-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206.269.0565
; TELEFAX: 206.269.0583
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 175 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-705-347A-53
Query Match 56.6%; Score 43; DB 3; Length 175;
Best Local Similarity 75.0%; Pred. No. 5.1;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GEEFSRAAEKLY 13
| | | | |
Db 18 GYEFSSRAAEALY 29
```

```

RESULT 8
US-09-324-542-53
; Sequence 53, Application US/09324542
; Patent No. 6328978
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L.J.
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Skin Disorders
; FILE REFERENCE: 11000.1007c1
; CURRENT APPLICATION NUMBER: US/09/324,542
; CURRENT FILING DATE: 1999-06-02
; EARLIER APPLICATION NUMBER: US 08/997,080
; EARLIER FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 53
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
US-09-324-542-53

```

```

Query Match      56.6%; Score 43; DB 3; Length 175;
Best Local Similarity 75.0%; Pred. No. 5.1;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      2 GEFSRAAEKLY 13
      |||||
DB      18 GYFSRACEALY 29

```

```

RESULT 9
US-09-205-426-53
; Sequence 53, Application US/09205426
; Patent No. 6406704
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L. J.
; TITLE OF INVENTION: Compounds and Methods for Treatment and
; TITLE OF INVENTION: Diagnosis of Mycobacterial Infections
; FILE REFERENCE: 11000.1002c4
; CURRENT APPLICATION NUMBER: US/09/205,426
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: 09/095,855
; EARLIER FILING DATE: 1998-06-11
; EARLIER APPLICATION NUMBER: 08/997,362
; EARLIER FILING DATE: 1997-12-23
; EARLIER APPLICATION NUMBER: 08/873,970
; EARLIER FILING DATE: 1997-06-12
; EARLIER APPLICATION NUMBER: 08/705,347
; EARLIER FILING DATE: 1996-08-29
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 53
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
US-09-205-426-53

```

```

Query Match      56.6%; Score 43; DB 4; Length 175;
Best Local Similarity 75.0%; Pred. No. 5.1;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      2 GEFSRAAEKLY 13
      |||||
DB      18 GYFSRACEALY 29

```

```

RESULT 10
US-09-200-643-53
; Sequence 53, Application US/09200643

```

```

; Patent No. 6410720
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L. J.
; TITLE OF INVENTION: Compounds and Methods for Treatment and
; TITLE OF INVENTION: Diagnosis of Mycobacterial Infections
; FILE REFERENCE: 11000.1002CON
; CURRENT APPLICATION NUMBER: US/09/200,643
; CURRENT FILING DATE: 1998-11-05
; EARLIER APPLICATION NUMBER: 08/705,347
; EARLIER FILING DATE: 1996-08-29
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 53
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
US-09-200-643-53

```

```

Query Match      56.6%; Score 43; DB 4; Length 175;
Best Local Similarity 75.0%; Pred. No. 5.1;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      2 GEFSRAAEKLY 13
      |||||
DB      18 GYFSRACEALY 29

```

```

RESULT 11
US-09-602-787A-314
; Sequence 314, Application US/09602787A
; Patent No. 6696561
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Mark
; APPLICANT: Kr"ger, Burkhard
; APPLICANT: Sch"der, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; TITLE OF INVENTION: TRANSPORT
; FILE REFERENCE: BGI-125CP
; CURRENT APPLICATION NUMBER: US/09/602,787A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: USN 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931454.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931478.0
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931563.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932122.1
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932124.8
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932180.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932182.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932190.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932191.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932209.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932212.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932227.9

```



```

; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932228.7
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932229.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932230.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932231.3
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933005.0
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933006.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19940764.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940765.7
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940766.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940830.0
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940831.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941395.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942077.7
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942078.5
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942079.3
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942088.2
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 678
; SEQ ID NO 314
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-602-787A-314

Query Match          56.6%; Score 43; DB 4; Length 708;
Best Local Similarity 53.3%; Pred. No. 23;
Matches      8; Conservative      3; Mismatches      4; Indels      0; Gaps      0;

QY      1 IGEEFSRAAEKLYLA 15
Db      360 VSEIAREVEKFLA 374

RESULT 12
US-09-602-787A-312
; Sequence 312, Application US/09602787A
; Patent No. 6696561
; GENERAL INFORMATION:
; APPLICANT: Pempejus, Mark
; APPLICANT: Kr"ger, Burkhard
; APPLICANT: Sch"der, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; FILE REFERENCE: EGI-125CP
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US/09/602,787A
; CURRENT FILING DATE: 2000-06-23
; SEQ ID NO 312
; LENGTH: 713
; PRIOR APPLICATION NUMBER: USSN 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931454.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931478.0
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931563.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932122.1
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932124.8
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932180.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932182.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932190.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932191.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932209.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932212.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932227.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932228.7
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932229.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932230.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932927.3
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933005.0
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933006.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19940764.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940765.7
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940766.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940830.0
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940831.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940832.7
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940833.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941395.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942077.7
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942078.5
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942079.3
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942088.2
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 678
; SEQ ID NO 312
; LENGTH: 713

```

```
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-602-787A-312

Query Match      56.6%; Score 43; DB 4; Length 713;
Best Local Similarity 53.3%; Pred. No. 23;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IGEEFSRAAEKLYLA 15
Db 365 VSBEIAREVEKFLA 379
      :||:|||||:|
      :||:|||||:|

RESULT 13
US-09-252-991A-18082
; Sequence 18082, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074.788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094.190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18082
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18082

Query Match      55.3%; Score 42; DB 4; Length 312;
Best Local Similarity 72.7%; Pred. No. 14;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FSRAAEKLYLA 15
Db 29 FSRAAEELFVA 39
      |||||:|:|
      |||||:|:|

RESULT 14
US-09-328-352-6612
; Sequence 6612, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328.352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6612
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6612

Query Match      55.3%; Score 42; DB 4; Length 324;
Best Local Similarity 53.8%; Pred. No. 15;
Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 3 EEFSRAAEKLYLA 15
Db 21 QSFTKAAEKLFIA 33
      :||:|||||:|
      :||:|||||:|

RESULT 15
US-09-538-092-1309
```

```
; Sequence 1309, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538.092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratSeqFormatter Version 0.9
; SEQ ID NO 1309
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number Q13573
US-09-538-092-1309

Query Match      55.3%; Score 42; DB 4; Length 536;
Best Local Similarity 53.3%; Pred. No. 26;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IGEEFSRAAEKLYLA 15
Db 280 INENFAKLAELALYA 294
      |||||:|:|
      |||||:|:|

Search completed: November 10, 2004, 15:57:23
Job time : 7.40214 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 02:43:24 ; Search time 38.6619 Seconds
(without alignments)
146.426 Million cell updates/sec

Title: US-10-092-750-225

Perfect score: 76

Sequence: 1 IGEFSPRAAEKLYLAV 16

Scoring table: BLOSUM62

Gap 10.0 , Gapext 0.5

Searched: 1568699 seqs, 353819137 residues

Total number of hits satisfying chosen parameters: 1568699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 76 | 100.0 | 16 | 14 | US-10-092-750-225 |
| 2 | 62 | 81.6 | 92 | 9 | US-09-925-300-1801 |
| 3 | 44 | 57.9 | 428 | 15 | US-10-403-571-110 |
| 4 | 44 | 57.9 | 876 | 9 | US-09-712-363-243 |
| 5 | 44 | 57.9 | 876 | 15 | US-10-282-122A-62407 |
| 6 | 44 | 57.9 | 876 | 15 | US-10-282-122A-64714 |
| 7 | 43 | 56.6 | 175 | 13 | US-09-880-505-53 |
| 8 | 43 | 56.6 | 175 | 13 | US-10-051-643-53 |
| 9 | 43 | 56.6 | 186 | 14 | US-10-369-493-23074 |
| 10 | 43 | 56.6 | 708 | 15 | US-10-627-476-314 |
| 11 | 43 | 56.6 | 713 | 15 | US-10-627-476-312 |
| 12 | 43 | 56.6 | 886 | 15 | US-10-282-122A-63869 |
| 13 | 43 | 56.6 | 1111 | 9 | US-09-738-626-6581 |

Sequence 444, App
Sequence 444, App
Sequence 61763, A
Sequence 86, Appl
Sequence 45692, A
Sequence 70, Appl
Sequence 63717, A
Sequence 37812, A
Sequence 8532, Ap
Sequence 179612, A
Sequence 132307, A
Sequence 68227, A
Sequence 1025, Ap
Sequence 1653, Ap
Sequence 1053, Ap
Sequence 636, App
Sequence 62077, A
Sequence 3642, Ap
Sequence 1982, Ap
Sequence 1983, Ap
Sequence 1984, Ap
Sequence 226931, A
Sequence 1078, Ap
Sequence 5, Appli
Sequence 354310, A
Sequence 11, Appl
Sequence 8311, Ap
Sequence 47758, A
Sequence 12979, A
Sequence 108934, A
Sequence 49573, A
Sequence 1285, Ap
Sequence 6501, Ap

323 9 US-09-925-302-444
323 10 US-09-925-302-444
323 15 US-10-282-122A-61763
323 15 US-10-403-571-86
323 15 US-10-282-122A-45692
323 14 US-10-166-225A-70
323 15 US-10-425-114-63717
323 16 US-10-767-701-37812
323 9 US-09-738-626-6532
323 16 US-10-437-963-179612
323 16 US-10-437-963-132307
323 15 US-10-282-122A-68227
323 15 US-10-389-566-1653
323 15 US-10-389-566-1025
323 15 US-10-389-566-636
323 15 US-10-282-122A-62077
323 14 US-10-369-493-3642
323 15 US-10-389-566-1982
323 15 US-10-389-566-1983
323 15 US-10-389-566-1984
323 15 US-10-424-599-226931
323 15 US-10-389-566-1078
323 9 US-09-796-149-5
323 17 US-10-425-115-354310
323 16 US-10-451-861-11
323 14 US-10-156-761-8311
323 15 US-10-282-122A-47758
323 14 US-10-156-761-12979
323 16 US-10-437-963-108934
323 15 US-10-282-122A-49573
323 9 US-09-925-300-1285
323 14 US-10-369-493-6501

ALIGNMENTS

RESULT 1
US-10-092-750-225
; Sequence 225, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hamond, Philip W.
; APPLICANT: Alpin, Julia
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 225
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-225

Query Match 100.0%; Score 76; DB 14; Length 16;
Best Local Similarity 100.0%; Pred. NO. 3.3e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IGEFSPRAAEKLYLAV 16
Db 1 IGEFSPRAAEKLYLAV 16

RESULT 2
US-09-925-300-1801
; Sequence 1801, Application US/09925300
; Patent No. US2002015181A1
; GENERAL INFORMATION:

```

; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1801
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (77)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (81)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (92)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1801

```

```

Query Match      81.6%; Score 62; DB 9; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.0059;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      3  EEFSRAAEKLYLA 15
Db      10  EEFSRAAEKLYLA 22

```

```

RESULT 3
US-10-403-571-110
; Sequence 110, Application US/10403571
; Publication No. US20040068763A1
; GENERAL INFORMATION:
; APPLICANT: Hopkins, Nancy
; APPLICANT: Golling, Gregory
; APPLICANT: Amsterdam, Adam
; APPLICANT: Sun, Zhaoxia
; TITLE OF INVENTION: Developmental Mutations in Zebrafish
; FILE REFERENCE: 01997/539002
; CURRENT APPLICATION NUMBER: US/10/403,571
; CURRENT FILING DATE: 2003-03-25
; PRIOR APPLICATION NUMBER: US 60/368,760
; PRIOR FILING DATE: 2002-03-29
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 110
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Danio rerio
US-10-403-571-110

```

```

Query Match      57.9%; Score 44; DB 15; Length 428;
Best Local Similarity 56.2%; Pred. No. 42;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```

```

QY      1  IGEEFSRAAEKLYLAV 16
Db      217  VGEYKLNQETLYLAV 232

```

```

RESULT 4
US-09-712-363-243
; Sequence 243, Application US/09712363
; Patent No. US20020164588A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Eisenberg, David
; APPLICANT: Rotstein, Sergio H.
; APPLICANT: Marcotte, Edward M.
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
; TITLE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
; FILE REFERENCE: 07419-032001
; CURRENT APPLICATION NUMBER: US/09/712,363
; CURRENT FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: PCT/US00/02246
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,531
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/117,844
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/118,206,
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/165,086
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 243
; LENGTH: 876
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-243

```

```

Query Match      57.8%; Score 44; DB 9; Length 876;
Best Local Similarity 75.0%; Pred. No. 90;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      2  GEFSRAAEKLY 13
Db      631  GYEFSSRACESLY 642

```

```

RESULT 5
US-10-282-122A-62407
; Sequence 62407, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Olesen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EUTRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06

```

```
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62407
; LENGTH: 876
; TYPE: PRT
; ORGANISM: Mycobacterium bovis
US-10-282-122A-62407

Query Match      57.9%; Score 44; DB 15; Length 876;
Best Local Similarity 75.0%; Pred. No. 90;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      2 GEEFSSRAAEKLY 13
Db      631 GYEFSSRACESLY 642

RESULT 6
US-10-282-122A-64714
; Sequence 64714, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64714
; LENGTH: 876
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-64714

Query Match      57.9%; Score 44; DB 15; Length 876;
Best Local Similarity 75.0%; Pred. No. 90;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      2 GEEFSSRAAEKLY 13
Db      631 GYEFSSRACESLY 642

RESULT 7
US-09-880-505-53
; Sequence 53, Application US/09880505
; Publication No. US20030007976A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L.J.
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Skin Disorders
; FILE REFERENCE: 11000.1007c2
; CURRENT APPLICATION NUMBER: US/09/880,505
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 09/324,542
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: US 08/997,080
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 53
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
US-09-880-505-53

Query Match      56.8%; Score 43; DB 10; Length 175;
Best Local Similarity 75.0%; Pred. No. 24;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      2 GEEFSSRAAEKLY 13
Db      18 GYEFSSRACEALY 29

RESULT 8
US-10-051-643-53
; Sequence 53, Application US/10051643
; Publication No. US20020197265A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L. J.
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Diseases of the Respiratory
; FILE REFERENCE: 11000.1008c2
; CURRENT APPLICATION NUMBER: US/10/051,643
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US09/156,181
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 08/996,624
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 53
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
```

US-10-051-643-53

Query Match 56.6%; Score 43; DB 13; Length 175;
 Best Local Similarity 75.0%; Pred. No. 24;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GEEFSRAAEKLY 13
 |||||
 Db 18 GYFESRACEALY 29

RESULT 9

US-10-369-493-23074
 ; Sequence 23074, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkie, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 23074
 ; LENGTH: 186
 ; TYPE: PRT
 ; ORGANISM: Bacillus subtilis
 US-10-369-493-23074

Query Match 56.6%; Score 43; DB 14; Length 186;
 Best Local Similarity 61.5%; Pred. No. 25;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 GEEFSRAAEKLYL 14
 |||||
 Db 62 GEDFSRKTEKEYI 74

RESULT 10

US-10-627-476-314
 ; Sequence 314, Application US/10627476
 ; Publication No. US20040030116A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pompeius, Mark
 ; APPLICANT: Kroger, Burkhard
 ; APPLICANT: Schoder, Hartwig
 ; APPLICANT: Zelder, Oskar
 ; APPLICANT: Habershauer, Gregor
 ; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
 ; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
 ; TITLE OF INVENTION: TRANSPORT
 ; FILE REFERENCE: BGI-125CPCN
 ; CURRENT APPLICATION NUMBER: US/10/627,476
 ; CURRENT FILING DATE: 2003-07-25
 ; PRIOR APPLICATION NUMBER: 09/602,787
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: USSN 60/141031
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: DE 19931454.3
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931478.0
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931563.9
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19932122.1
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: DE 19932180.9
 ; PRIOR FILING DATE: 1999-07-09
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 678
 ; SEQ ID NO 312
 ; LENGTH: 713
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-10-627-476-312

Query Match 56.6%; Score 43; DB 15; Length 713;
 Best Local Similarity 53.3%; Pred. No. 11e+02;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 IGEEFSRAAEKLYLA 15

; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: DE 19932125.6
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: DE 19932128.0
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: DE 19932180.9
 ; PRIOR FILING DATE: 1999-07-09
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 678
 ; SEQ ID NO 314
 ; LENGTH: 708
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-10-627-476-314

Query Match 56.6%; Score 43; DB 15; Length 708;
 Best Local Similarity 53.3%; Pred. No. 11e+02;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 IGEEFSRAAEKLYLA 15
 :|||:|||||
 Db 360 VSEBIAREVEKFLA 374

RESULT 11

US-10-627-476-312
 ; Sequence 312, Application US/10627476
 ; Publication No. US20040030116A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pompeius, Mark
 ; APPLICANT: Kroger, Burkhard
 ; APPLICANT: Schoder, Hartwig
 ; APPLICANT: Zelder, Oskar
 ; APPLICANT: Habershauer, Gregor
 ; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
 ; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
 ; TITLE OF INVENTION: TRANSPORT
 ; FILE REFERENCE: BGI-125CPCN
 ; CURRENT APPLICATION NUMBER: US/10/627,476
 ; CURRENT FILING DATE: 2003-07-25
 ; PRIOR APPLICATION NUMBER: 09/602,787
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: USSN 60/141031
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: DE 19931454.3
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931478.0
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931563.9
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19932122.1
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: DE 19932124.8
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: DE 19932125.6
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: DE 19932128.0
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: DE 19932180.9
 ; PRIOR FILING DATE: 1999-07-09
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 678
 ; SEQ ID NO 312
 ; LENGTH: 713
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-10-627-476-312

Query Match 56.6%; Score 43; DB 15; Length 713;
 Best Local Similarity 53.3%; Pred. No. 11e+02;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 IGEEFSRAAEKLYLA 15

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:52:32, Search time 5.46619 Seconds
(without alignments)
281.634 Million cell updates/sec

Title: US-10-092-750-225

Perfect score: 76

Sequence: 1 IGEEFSRAAEKLYLA 16

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 62 | 81.6 | 86 | 2 A57292 | signal recognition |
| 2 | 62 | 81.6 | 86 | 2 A34731 | signal recognition |
| 3 | 62 | 81.6 | 86 | 2 S57500 | signal recognition |
| 4 | 44 | 57.9 | 403 | 2 B72778 | probable processin |
| 5 | 44 | 57.9 | 415 | 2 I51637 | cyclin A2 - Africa |
| 6 | 44 | 57.9 | 876 | 2 G70863 | probable vals prot |
| 7 | 43 | 56.6 | 56 | 2 S53000 | mitotic-specific c |
| 8 | 43 | 56.6 | 186 | 1 I39782 | shikimate kinase (|
| 9 | 43 | 56.6 | 547 | 2 A56575 | puff-specific nucl |
| 10 | 43 | 56.6 | 886 | 2 A87093 | valyl-trna synthas |
| 11 | 42 | 55.3 | 242 | 2 T32122 | hypothetical prote |
| 12 | 42 | 55.3 | 302 | 2 F83174 | probable transcrip |
| 13 | 42 | 55.3 | 418 | 2 S11678 | cyclin A - African |
| 14 | 41 | 53.9 | 191 | 1 H69472 | iron-dependent rep |
| 15 | 41 | 53.9 | 312 | 2 B97313 | transcription regu |
| 16 | 41 | 53.9 | 327 | 2 H96512 | probable cyclin, 2 |
| 17 | 41 | 53.9 | 340 | 2 T19538 | hypothetical prote |
| 18 | 41 | 53.9 | 341 | 2 T51897 | related to sorbito |
| 19 | 41 | 53.9 | 349 | 2 H69162 | conserved hypothet |
| 20 | 41 | 53.9 | 465 | 2 S76164 | hypothetical prote |
| 21 | 40 | 52.6 | 227 | 2 S71193 | mitosis-specific c |
| 22 | 40 | 52.6 | 295 | 2 AE1019 | lysr family regula |
| 23 | 40 | 52.6 | 445 | 2 B71642 | hypothetical prote |
| 24 | 40 | 52.6 | 452 | 2 F86289 | probable cyclin [i |
| 25 | 40 | 52.6 | 454 | 2 A96803 | probable mitotic c |
| 26 | 40 | 52.6 | 882 | 2 S41034 | hypothetical prote |
| 27 | 39 | 51.3 | 76 | 2 S40765 | hypothetical prote |
| 28 | 39 | 51.3 | 96 | 2 H83065 | hypothetical prote |
| 29 | 39 | 51.3 | 133 | 2 F65179 | psa3 protein - Esc |

30 39 51.3 195 2 G95912 probable transposase
31 39 51.3 279 2 B92166 regulator of bssa
32 39 51.3 293 2 AF0475 lysr-family transcr
33 39 51.3 301 2 B83416 probable transcrip
34 39 51.3 306 2 C82387 transcripction regu
35 39 51.3 308 2 D72768 probable electron
36 39 51.3 316 2 S75062 transcripction regu
37 39 51.3 328 2 AB0435 conserved hypothet
38 39 51.3 395 2 S36812 cyclin A - chicken
39 39 51.3 406 2 S24788 cyclin A - bovine
40 39 51.3 422 2 S37280 cyclin A2 - mouse
41 39 51.3 422 2 S38501 cyclin A - human
42 39 51.3 432 2 S08277 Similar to Cytochr
43 39 51.3 476 2 D86306 hypothetical prote
44 39 51.3 535 2 T25379 hypothetical prote
45 39 51.3 539 2 H72422

ALIGNMENTS

RESULT 1

A57292 signal recognition particle protein SRP9 - human

C;Species: Homo sapiens (man)

C;Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 09-Jul-2004

C;Accession: A57292

R;Hsu, K.; Chang, D.Y.; Maraiia, R.J.

J. Biol. Chem. 270, 10179-10186, 1995

A;Title: Human signal recognition particle (SRP) Alu-associated protein also binds Alu ir

A;Reference number: A57292; MUID:95247726; PMID:7730321

A;Accession: A57292

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-86 <HSU>

A;Cross-references: UNIPROT:P49458; GB:U20998; NID:G897850; PIDN:AAA70170.1; PID:G897851

A;Note: authors translated the codon CCG for residue 2 as Ala, and GAG for residue 86 as

C;Genetics:

A;Gene: GDB:SRP9

A;Cross-references: GDB:128862; OMIM:600707

A;Map position: 5q21-5q22

Query Match 81.6%; Score 62; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.00089; Mismatches 0; Indels 0; Gaps 0;
Matches 13; Conservative 0

Qy 3 EEFSRAAEKLYLA 15

Db 8 EEFSRAAEKLYLA 20

RESULT 2

A34731 signal recognition particle 9K chain - dog

C;Species: Canis lupus familiaris (dog)

C;Date: 06-Jul-1990 #sequence_revision 06-Jul-1990 #text_change 09-Jul-2004

C;Accession: A34731

R;Strub, K.; Walter, P.

Mol. Cell. Biol. 10, 777-784, 1990

A;Title: Assembly of the Alu domain of the signal recognition particle (SRP): dimerizatio

A;Reference number: A34731; MUID:90136594; PMID:2153922

A;Accession: A34731

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-86 <STR>

A;Cross-references: UNIPROT:P21262; GB:M34952; NID:G164085; PIDN:AAA30897.1; PID:G164086

Query Match 81.6%; Score 62; DB 2; Length 86;

Best Local Similarity 100.0%; Pred. No. 0.00089;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EEFSRAAEKLYLA 15

Db 8 EEFSRAAEKLYLA 20

Db 8 EEF5RAAEKLYLA 20

RESULT 3
S57500
signal recognition particle - mouse
C:Species: Mus musculus (house mouse)
C>Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: S57500
R;Bui, N.; Bova, F.; Wolff, N.; Morrical, M.D.; Walter, P.; Strub, K.
A:Submitted to the EMBL Data Library, March 1994
A:Description: The 9kD subunit of the murine signal recognition particle.
A:Reference number: S57500
A:Accession: S57500
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-86 <BUI>
A:Cross-references: UNIPROT:P49962; EMBL:X78304; NID:g8721130; PIDN:CAA55114.1; PID:g8721130

Query Match 81.6%; Score 62; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.00089;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EEF5RAAEKLYLA 15
:|||||
Db 8 EEF5RAAEKLYLA 20
:|||||

RESULT 4
B72778
probable processing proteinase APE0212 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: B72778
R;Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; K DNA Res. 5, 83-101, 1999
A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: B72778
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-403 <RAW>
A:Cross-references: UNIPROT:Q9YFN7; DDBJ:AP000058; NID:95103388; PIDN:BAA79124.1; PID:g872778
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE0212
C:Superfamily: mitochondrial processing peptidase alpha chain

Query Match 57.9%; Score 44; DB 2; Length 403;
Best Local Similarity 50.0%; Pred. No. 7;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IGEEFSRAAEKLYLA 16
:|||||
Db 93 VSDSLARVAEKLYLA 108
:|||||

RESULT 5
I51637
cyclin A2 - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 12-Jul-2004
C:Accession: I51637; S52713
R;Howe, J.A.; Howell, M.; Hunt, T.; Newport, J.W.
Genes Dev. 9, 1164-1176, 1995
A>Title: Identification of a developmental timer regulating the stability of embryonic cyclin A
A:Reference number: I51637; MUID:95278730; PMID:7758942
A:Accession: I51637
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-415 <HOW>
A:Cross-references: UNIPROT:P47827; EMBL:X85746; NID:g755807; PIDN:CAA55748.1; PID:g755807

C:Superfamily: cyclin, A/B/D/E type
C:Keywords: cell cycle control

Query Match 57.9%; Score 44; DB 2; Length 415;
Best Local Similarity 56.2%; Pred. No. 7.2;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IGEEFSRAAEKLYLA 16
:|||||
Db 203 VGEEYKLNQETLYLA 218
:|||||

RESULT 6
G70863
Probable vals protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: G70863
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Ruter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrrell, B.G.
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: G70863
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-876 <COL>
A:Cross-references: UNIPROT:O53175; GB:AL021246; GB:AL123456; NID:g3261507; PIDN:CAA16021
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: vals
C:Superfamily: valine-tRNA ligase

Query Match 57.9%; Score 44; DB 2; Length 876;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GEEFSRAAEKLY 13
:|||||
Db 631 GYEFSPACESLY 642
:|||||

RESULT 7
S53000
mitotic-specific cyclin CYC2.3 - rape (fragment)
C:Species: Brassica napus (rape)
C>Date: 21-Apr-1997 #sequence_revision 21-Apr-1997 #text_change 12-Jul-2004
C:Accession: S53000
R;Szarka, S.; Fitch, M.; Schaerer, S.; Moloney, M.
Plant Mol. Biol. 27, 263-275, 1995
A>Title: Classification and expression of a family of cyclin gene homologues in Brassica
A:Reference number: S52996; MUID:95195155; PMID:788617
A:Accession: S53000
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-56 <SZA>
A:Cross-references: UNIPROT:Q39327; EMBL:L25402
A>Note: the sequence of residues 1-2, 51-56 and the corresponding nucleic acid sequence
C:Genetics:
A:Introns: 5/3; 50/2
C:Superfamily: cyclin, A/B/D/E type
C:Keywords: cell cycle control; cell division; mitosis

Query Match 56.6%; Score 43; DB 2; Length 56;
Best Local Similarity 50.0%; Pred. No. 1.4;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IGEEFSRAAEKLYLA 16
:|||||
Db 6 VSEYKIASDTLYLA 21
:|||||

```

Query Match          56.6%;   Score 43; DB 2; Length 547;
Best Local Similarity 53.3%;   Pred. No. 15;
Matches            8; Conservative    4; Mismatches    3; Indels    0; Gaps    0;

QY      1 IGEEFSRAAEKLYLA 15
|||::|||::|||
DB      283 INEPAKWAELIYA 297

RESULT 10
AB87093
valyl-tRNA synthase [imported] - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: AB87093
R;Cole,S.T.; Eglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Horrold, R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd, C.; Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Stewardson, K.; Wainwright, M.A.; Rutherford, K.M. Nature 409, 1007-1011, 2001
A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Stewardson, K.; Wainwright, M.A.; Rutherford, K.M. Nature 409, 1007-1011, 2001
A;Title: Massive gene decay in the leprosy bacillus.
A;Reference number: AB86909; MUID:21128732; PMID:11234002
A;Accession: AB87093
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-886 <STO>
A;Cross-references: UNIPROT:Q9CBY7; GB:AL450380; NID:g13093326; PIDN:CAC30422.1; GSPDB:GN87093
C;Genetics:
A;Gene: vals
C;Superfamily: valine-tRNA ligase

Query Match          56.6%;   Score 43; DB 2; Length 886;
Best Local Similarity 75.0%;   Pred. No. 24;
Matches            9; Conservative    0; Mismatches    3; Indels    0; Gaps    0;

QY      2 GEESFRAAEKLY 13
|||||::|||
DB      641 GYEFSFACEALY 652

RESULT 11
T32122
hypothetical protein F59E11.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T32122
R;Bradshaw, H. submitted to the EMBL Data Library, July 1997
A;Description: The sequence of C. elegans cosmid F59E11.
A;Reference number: Z21124
A;Accession: T32122
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-242 <BRA>
A;Cross-references: UNIPROT:O16761; EMBL:AF016685; PIDN:AAB66219.1; GSPDB:GN00023; CESP:F83174
A;Experimental source: strain Bristol N2; clone F59E11
C;Genetics:
A;Gene: CESP:F59E11.5
A;Map position: 5
A;Introns: 87/3; 123/3; 180/3; 217/3

Query Match          55.3%;   Score 42; DB 2; Length 242;
Best Local Similarity 66.7%;   Pred. No. 9.3;
Matches            8; Conservative    2; Mismatches    2; Indels    0; Gaps    0;

QY      3 EEFSRAAEKLYL 14
:|||::|||::
DB      152 KEFSCEAKLYV 163

RESULT 12
F83174
```

probable transcription regulator PA3776 [imported] - Pseudomonas aeruginosa (strain PA01)
 C:Species: Pseudomonas aeruginosa
 C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
 C/Accession: F83174
 R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B.
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Latbig, K.; Lim,
 ; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
 A/Reference number: A82950; MUID:20437337; PMID:10984043
 A/Accession: F83174
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-302 <STO>
 A/Cross-references: UNIPROT:Q9HXL9; GB:AE004796; GB:AE004091; NID:g9949939; PIDN:AAG0716
 A/Experimental source: strain PA01
 C/Genetics:
 A/Gene: PA3776
 C/Superfamily: Pseudomonas putida regulatory protein caSR

Query Match 55.3%; Score 42; DB 2; Length 302;
 Best Local Similarity 72.7%; Pred. No. 12;
 Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 5 FSPAAEKLYLA 15
 |||||:|:|:
 Db 19 FSPAAERLFA 29
 |||||:|:|:

RESULT 13
 S11678
 Cyclin A - African clawed frog
 C/Species: Xenopus laevis (African clawed frog)
 C/Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 12-Jul-2004
 C/Accession: S11678
 R/Minshull, J.; Golsteyn, R.; Hill, C.S.; Hunt, T.
 EMBO J. 9, 2865-2875, 1990
 A>Title: The A- and B-type cyclin associated cdc2 kinases in Xenopus turn on and off at
 A/Reference number: S11678; MUID:90360999; PMID:2143393
 A/Accession: S11678
 A/Molecule type: mRNA
 A/Residues: 1-418 <MIN>
 A/Cross-references: UNIPROT:P18606; EMBL:X53745; NID:g64644; PIDN:CAA37775.1; PID:g64645
 C/Superfamily: cyclin, A/B/D/E type
 C/Keywords: cell cycle control

Query Match 55.3%; Score 42; DB 2; Length 418;
 Best Local Similarity 50.0%; Pred. No. 17;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 IGEEFSPRAAEKLYLA 16
 |||||:|:|:
 Db 207 VGEEYKLHTETLYLA 222
 |||||:|:|:

RESULT 14
 H69472
 iron-dependent repressor homolog - Archaeoglobus fulgidus
 C/Species: Archaeoglobus fulgidus
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C/Accession: H69472
 R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A/Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A>Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
 A/Reference number: A69250; MUID:98049343; PMID:9389475
 A/Accession: H69472
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-191 <KLE>

A/Cross-references: UNIPROT:O28489; GB:AE000979; GB:AE000782; NID:g2689302; PIDN:AAB89462
 C/Superfamily: conserved hypothetical protein MJ0568

Query Match 53.9%; Score 41; DB 1; Length 191;
 Best Local Similarity 53.3%; Pred. No. 11;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 GEEPSRAAEKLYLA 16
 |||||:|:|:
 Db 65 GEELARRIKKYLAL 79
 |||||:|:|:

RESULT 15
 B97313
 transcription regulator, LysR family [imported] - Clostridium acetobutylicum
 C/Species: Clostridium acetobutylicum
 C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
 C/Accession: B97313
 R/Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
 ; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4823-4838, 2001
 A>Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
 A/Reference number: A96900; MUID:21359325; PMID:21359325
 A/Accession: B97313
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-312 <KUR>
 A/Cross-references: UNIPROT:Q97DV9; GB:AE001437; PIDN:AAK81293.1; PID:g15026445; GSPDB:G
 A/Experimental source: Clostridium acetobutylicum ATCC824
 C/Genetics:
 A/Gene: CAC3361

Query Match 53.9%; Score 41; DB 2; Length 312;
 Best Local Similarity 58.8%; Pred. No. 18;
 Matches 10; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 1 IGEE--PSRAAEKLYLA 15
 |||||:|:|:
 Db 12 IAEKSISRRAERLYIS 28
 |||||:|:|:

Search completed: November 10, 2004, 15:55:08
 Job time : 5.46619 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:50:40 ; Search time 26.8185 Seconds
(without alignments)
343.270 Million cell updates/sec

Title: US-10-092-750-225
Perfect score: 76
Sequence: 1 IGEEFSRAAEKLYLA 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 02: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------------|----------------------|
| 1 | 62 | 81.6 | 49 | 2 Q6P2S0 | Q6P2S0 homo sapien |
| 2 | 62 | 81.6 | 49 | 2 AAh64351 | AAh64351 homo sapi |
| 3 | 62 | 81.6 | 82 | 2 Q6VUX0 | Q6VUX0 homo sapien |
| 4 | 62 | 81.6 | 82 | 2 Q8WTW0 | Q8WTW0 homo sapien |
| 5 | 62 | 81.6 | 82 | 2 AAh67845 | AAh67845 homo sapi |
| 6 | 62 | 81.6 | 85 | 1 SR09 CANFA | P21262 canis famil |
| 7 | 62 | 81.6 | 85 | 1 SR09 HUMAN | P49458 homo sapien |
| 8 | 62 | 81.6 | 85 | 1 SR09 MOUSE | P49962 mus musculu |
| 9 | 62 | 81.6 | 86 | 2 AAh65957 | AAh65957 homo sapi |
| 10 | 62 | 81.6 | 86 | 2 BAC34557 | BAC34557 mus muscu |
| 11 | 55 | 72.4 | 86 | 2 Q7ZW94 | Q7ZW94 brachydania |
| 12 | 50 | 65.8 | 464 | 2 Q8IY91 | Q8IY91 homo sapien |
| 13 | 50 | 65.8 | 465 | 1 CGA1 HUMAN | P78396 homo sapien |
| 14 | 48 | 63.2 | 390 | 2 Q7T3L6 | Q7T3L6 brachydania |
| 15 | 48 | 63.2 | 395 | 2 Q8G3P3 | Q8G3P3 bifidobacte |
| 16 | 47 | 61.8 | 301 | 2 Q7NG14 | Q7NG14 gloeobacter |
| 17 | 46 | 60.5 | 421 | 1 CGA1 MOUSE | Q61456 mus musculu |
| 18 | 46 | 60.5 | 421 | 2 Q8C5U1 | Q8C5U1 mus musculu |
| 19 | 44 | 57.9 | 270 | 2 Q785S2 | Q785S2 klebsiella |
| 20 | 44 | 57.9 | 270 | 2 BAD14988 | BAD14988 klebsiell |
| 21 | 44 | 57.9 | 391 | 1 CGA1 CARAU | P21161 carassius a |
| 22 | 44 | 57.9 | 391 | 2 Q9PW42 | Q9PW42 carassius a |
| 23 | 44 | 57.9 | 391 | 2 Q9PW43 | Q9PW43 carassius a |
| 24 | 44 | 57.9 | 403 | 2 Q9YFN7 | Q9YFN7 aeropyrum p |
| 25 | 44 | 57.9 | 410 | 2 Q6NV43 | Q6NV43 brachydania |
| 26 | 44 | 57.9 | 410 | 2 Q7ZVJ2 | Q7ZVJ2 brachydania |
| 27 | 44 | 57.9 | 410 | 2 AAh68323 | AAh68323 brachydania |
| 28 | 44 | 57.9 | 415 | 1 CGA2 XENLA | P47827 xenopus lae |
| 29 | 44 | 57.9 | 415 | 2 Q6DB50 | Q6DB50 xenopus lae |
| 30 | 44 | 57.9 | 415 | 2 Q6DIH8 | Q6DIH8 xenopus tro |
| 31 | 44 | 57.9 | 428 | 2 Q918J1 | Q918J1 carassius a |

32 44 57.9 428 2 Q98TA3 Q98ta3 brachydania
33 44 57.9 536 2 Q8J2Z0 Q8j2z0 methanococc
34 44 57.9 536 2 CAF30163 Caf30163 methanoco
35 44 57.9 876 1 SYV MYCTU O53175 mycobacteri
36 43 56.6 57 2 Q39327 Q39327 brassica na
37 43 56.6 186 1 AROK BACSU F37944 bacillus su
38 43 56.6 289 2 Q8XHM1 Q8xhm1 clostridium
39 43 56.6 301 2 Q7VV47 Q7vv47 bordetella
40 43 56.6 307 2 Q92JX0 Q92jx0 rhizobium m
41 43 56.6 329 2 Q7Q9Y9 Q7q9y9 anopheles g
42 43 56.6 344 2 Q9RDL1 Q9rdl1 streptomyce
43 43 56.6 393 2 Q9VDD0 Q9vdd0 drosophila
44 43 56.6 393 2 AAQ23542 Aaq23542 drosophila
45 43 56.6 547 1 BX42_DROME F39736 drosophila

ALIGNMENTS

RESULT 1
Q6P2S0 PRELIMINARY; PRT; 49 AA.
ID Q6P2S0
AC Q6P2S0;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE SRP9 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Narusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.D., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
FAhey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalish D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX Submitted (JSC-2003) to the EMBL/GenBank/DBJ databases.
RA Strausberg R.;
RL EMBL; EC064351; AAh64351.1; -;
DR InterPro; IPR008832; SRP9.
DR InterPro; IPR009018; SRP9/14.
DR Pfam; PF05486; SRP9; 1.
SQ SEQUENCE 49 AA; 5769 MW; 3E90E32103B227DF CRC64;

Query Match 81.6%; Score 52; DB 2; Length 49;
Best Local Similarity 100.0%; Pred. No. 0.0036;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EEFSSRAAEKLYLA 15
| | | | | | | | | | | | | | | |
Db 8 EEFSSRAAEKLYLA 20

```

RA  AAH64351      PRELIMINARY;          PRT;          49 AA.
RESULT 2
AC  AAH64351;
DT  02-MAR-2004 (TEMBLrel. 27, Created)
DT  02-MAR-2004 (TEMBLrel. 27, Last sequence update)
DT  02-MAR-2004 (TEMBLrel. 27, Last annotation update)
DE  SRP9 protein.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  MEDLINE=2238257; PubMed=12477932;
RA  Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA  Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA  Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA  Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA  Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA  Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA  Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA  Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA  Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA  Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA  Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA  Whiting M., Touchman J.W., Green E.D., Dickinson M.C.,
RA  Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA  Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA  Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA  Jones S.J., Marra M.A.;
RT  "Generation and initial analysis of more than 15,000 full-length human
RL  and mouse cDNA sequences.";
RP  Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Kidney;
RA  Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA  Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA  Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA  Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA  Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA  Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA  Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA  Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA  Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA  Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA  Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA  Whiting M., Touchman J.W., Green E.D., Dickinson M.C.,
RA  Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA  Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA  Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA  Jones S.J., Marra M.A.;
RT  "Generation and initial analysis of more than 15,000 full-length human
RL  and mouse cDNA sequences.";
RP  Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Kidney;
RA  Strausberg R.;
RA  Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR  EMBL; BC064351; AAH64351.1; -
DR  InterPro; IPR009018; SRP9/14.
DR  Pfam; PF05486; SRP9; 1.
SQ  SEQUENCE 49 AA; 5769 MW; 3E90E32103B227DF CRC64;

Query Match          81.6%; Score 62; DB 2; Length 49;
Best Local Similarity 100.0%; Pred. No. 0.0036;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  3 EEF5RAAEKLYLA 15
DB  8 EEF5RAAEKLYLA 20
|||||
|||||

RESULT 3
Q6NVX0      PRELIMINARY;          PRT;          82 AA.
AC  Q6NVX0;
DT  05-JUL-2004 (TEMBLrel. 27, Created)
DT  05-JUL-2004 (TEMBLrel. 27, Last sequence update)
DT  05-JUL-2004 (TEMBLrel. 27, Last annotation update)
DE  SRP9 protein.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Brain;
RA  Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA  Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA  Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA  Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA  Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA  Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA  Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA  Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA  Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA  Villallon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA  Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA  Whiting M., Touchman J.W., Green E.D., Dickinson M.C.,
RA  Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA  Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA  Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA  Jones S.J., Marra M.A.;
RT  "Generation and initial analysis of more than 15,000 full-length human
RL  and mouse cDNA sequences.";
RP  Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Kidney;
RA  Strausberg R.;
RA  Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR  EMBL; BC064351; AAH64351.1; -
DR  InterPro; IPR009018; SRP9/14.
DR  Pfam; PF05486; SRP9; 1.
SQ  SEQUENCE 49 AA; 5769 MW; 3E90E32103B227DF CRC64;

Query Match          81.6%; Score 62; DB 2; Length 49;
Best Local Similarity 100.0%; Pred. No. 0.0036;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  3 EEF5RAAEKLYLA 15
DB  8 EEF5RAAEKLYLA 20
|||||
|||||

RESULT 4
Q8WTW0      PRELIMINARY;          PRT;          82 AA.
AC  Q8WTW0;
DT  01-MAR-2002 (TEMBLrel. 20, Created)
DT  01-MAR-2002 (TEMBLrel. 20, Last sequence update)
DT  01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE  SRP9 protein.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Bone marrow;
RA  Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA  Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA  Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA  Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA  Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA  Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA  Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA  Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA  Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA  Villallon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA  Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA  Whiting M., Touchman J.W., Green E.D., Dickinson M.C.,
RA  Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA  Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA  Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA  Jones S.J., Marra M.A.;
RT  "Generation and initial analysis of more than 15,000 full-length human
RL  and mouse cDNA sequences.";
RP  Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Brain;
RA  Strausberg R.;
RA  Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR  EMBL; BC067845; AAH67845.1; -
DR  InterPro; IPR008832; SRP9
DR  InterPro; IPR009018; SRP9/14.
DR  Pfam; PF05486; SRP9; 1.
SQ  SEQUENCE 82 AA; 9124 MW; D4C817FE1A6D7ED5 CRC64;

Query Match          81.6%; Score 62; DB 2; Length 82;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  3 EEF5RAAEKLYLA 15
DB  8 EEF5RAAEKLYLA 20
|||||
|||||

```

```

RT and mouse cDNA sequences.":
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC021995; AAH21995.1; -.
DR HSSP; P49458; 1E80.
DR GO; GO:0005786; C:signal recognition particle; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0045900; P:negative regulation of translational elonga. .; IEA.
DR GO; GO:0006605; P:protein targeting; IEA.
DR InterPro; IPR008832; SRP9.
DR InterPro; IPR009018; SRP9/14.
DR Pfam; PF05486; SRP9; 1.
SQ SEQUENCE 82 AA; 9142 MW; D4C81093DA6D7ED5 CRC64;

Query Match 81.6%; Score 62; DB 2; Length 82;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EEF5RAAEKLYLA 15
DB 8 EEF5RAAEKLYLA 20

RESULT 5
AAH67845 PRELIMINARY; PRT; 82 AA.
ID AAH67845
AC AAH67845
DT 14-APR-2004 (TrEMBLrel. 27, Created)
DT 14-APR-2004 (TrEMBLrel. 27, Last sequence update)
DE SRP9 protein.
GN SRP9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Teshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A.C., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC067845; AAH67845.1; -.
SQ SEQUENCE 82 AA; 9124 MW; D4C817FE1A6D7ED5 CRC64;

Query Match 81.6%; Score 62; DB 2; Length 82;

Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EEF5RAAEKLYLA 15
DB 8 EEF5RAAEKLYLA 20

RESULT 6
SR09 CANFA STANDARD; PRT; 85 AA.
ID SR09 CANFA
AC P21262;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Signal recognition particle 9 kDa protein (SRP9).
GN Name=SRP9;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-31.
RX MEDLINE=90136594; PubMed=2153922;
RA Strub K., Walter P.;
RT "Assembly of the Aiu domain of the signal recognition particle (SRP):
RT dimerization of the two protein components is required for efficient
RT binding to SRP RNA."
RL Mol. Cell. Biol. 10:777-784(1990).
CC -!- FUNCTION: Signal-recognition-particle assembly has a crucial role
CC in targeting secretory proteins to the rough endoplasmic reticulum
CC membrane. SRP9 together with SRP14 and the Aiu portion of the SRP
CC RNA, constitutes the elongation arrest domain of SRP. The complex
CC of SRP9 and SRP14 is required for SRP RNA binding.
CC -!- SUBUNIT: Signal recognition particle consists of a 7S RNA molecule
CC of 300 nucleotides and six protein subunits: SRP72, SRP68, SRP54,
CC SRP19, SRP14 and SRP9.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the SRP9 family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; M34952; AAA30897.1; -.
DR PIR; A34731; A34731.
DR HSSP; P49458; 1E80.
DR InterPro; IPR008832; SRP9.
DR InterPro; IPR009018; SRP9/14.
DR Pfam; PF05486; SRP9; 1.
KW Direct protein sequencing; RNA-binding; Signal recognition particle.
FT INIT MET
SQ SEQUENCE 85 AA; 9914 MW; 047EA875B270B450 CRC64;

Query Match 81.6%; Score 62; DB 1; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.0063;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EEF5RAAEKLYLA 15
DB 7 EEF5RAAEKLYLA 19

RESULT 7
SR09 HUMAN STANDARD; PRT; 85 AA.
ID SR09 HUMAN
AC P49458;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)

```

01-OCT-2004 (Rel. 45, Last annotation update)
 Signal recognition particle 9 kDa protein (SRP9).

GN Name=SRP9;
 OS Homo sapiens (Human);
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN
 RP
 RX MEDLINE=95247726; PubMed=7730321;
 RA Hsu K., Chang D.-Y., Marais R.J.;
 RT "Human signal recognition particle (SRP) Alu-associated protein also
 binds Alu interspersed repeat sequence RNAs. Characterization of human
 SRP9.";
 RL J. Biol. Chem. 270:10179-10186(1995).
 [2]
 RN
 RP
 RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S.E., Garcia A.M., Gay L.J., Huiyk S.W.,
 RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J.J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shcherchenko Y., Bouffard G.G.,
 RA Blackley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC
 CC -1- FUNCTION: Signal-recognition-particle assembly has a crucial role
 in targeting secretory proteins to the rough endoplasmic reticulum
 membrane. SRP9 together with SRP14 and the Alu portion of the SRP
 RNA, constitutes the elongation arrest domain of SRP. The complex
 of SRP9 and SRP14 is required for SRP RNA binding.
 CC
 CC -1- SUBUNIT: Signal recognition particle consists of a 7S RNA molecule
 of 300 nucleotides and six protein subunits: SRP72, SRP68, SRP54,
 SRP19, SRP14 and SRP9.
 CC
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC
 CC -1- SIMILARITY: Belongs to the SRP9 family.
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL: U20998; AA070170.1; -;
 CC EMBL: BC008443; AA08443.1; -;
 CC EMBL: BC015094; AA015094.1; -;
 CC EMBL: BC022415; AA022415.1; -;
 CC PIR: A57292; A57292.
 CC PDB: 1E80; X-ray; A/Cs1-85.
 CC PDB: 1E85; X-ray; A=1-85.
 CC InterAct: P49458; -;
 CC Genew: HGNC:11304; SRP9.
 CC MIM: 600707; -;
 CC GO: GO:0005785; C:signal recognition particle receptor complex; TAS.
 CC GO: GO:0003723; F:RNA binding; TAS.
 CC GO: GO:0005047; F:signal recognition particle binding; TAS.
 CC InterPro: IPR008832; SRP9.
 CC InterPro: IPR009018; SRP9/14.

DR Pfam; PF05486; SRP9; 1.
 KW 3D-structure; RNA-binding; Signal recognition particle.
 FT INIT MET 0 0 By similarity.
 FT STRAND 2 3
 FT HELIX 6 18
 FT TURN 19 19
 FT TURN 21 23
 FT STRAND 25 31
 FT HELIX 32 34
 FT TURN 35 35
 FT STRAND 36 42
 FT TURN 47 52
 FT STRAND 55 55
 FT TURN 56 71
 FT TURN 72 73
 SQ SEQUENCE 85 AA; 9980 MW; BFB1E9ECAF70A0E4 CRC64;
 Query Match 81.6%; Score 62; DB 1; Length 85;
 Best Local Similarity 100.0%; Pred. NO. 0.0063;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 EEFSSRAAEKLYLA 15
 DQ |||||
 ID SR09 MOUSE STANDARD; PRT; 85 AA.
 AC P49962; O9D085;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Signal recognition particle 9 kDa protein (SRP9).
 GN Name=SRP9;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP
 RX MEDLINE=94301784; PubMed=7518078;
 RA Bova F., Bui N., Strub K.;
 RT "The heterodimeric subunit SRP9/14 of the signal recognition particle
 functions as permuted single polypeptide chain";
 RL Nucleic Acids Res. 22:2028-2035(1994).
 [2]
 RN
 RP
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Sakakawa T., Adachi J., Boro H., Kondo S.,
 RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojbori T.,
 RA Badarrelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schirini L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedziera R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Perle G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming D.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imctani K., Ishii Y., Itoh M., Kagawa I.,

Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Birney E., Hayashizaki Y.;
RA "Analysis of the mouse transcriptome based on functional annotation of
RA 60,770 full-length cDNAs";
RT Nature 420:563-573(2002).
RL [3]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary gland;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Mariani K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RA "generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP RAY CRYSTALLOGRAPHY (2.53 ANGSTROMS) IN COMPLEX WITH SRP14.
RX MEDLINE=97377014; PubMed=9233785;
RA Birse D.E., Kapp U., Strub K., Cusack S., Aaberg A.;
RA "The crystal structure of the signal recognition particle Alu RNA
RT binding heterodimer, SRP9/14";
RT EMBO J. 16:3757-3766(1997).
CC -|- FUNCTION: Signal-recognition-particle assembly has a crucial role
CC in targeting secretory proteins to the rough endoplasmic reticulum
CC membrane. SRP9 together with SRP14 and the Alu portion of the SRP
CC RNA, constitutes the elongation arrest domain of SRP. The complex
CC of SRP9 and SRP14 is required for SRP RNA binding.
CC -|- SUBUNIT: Signal recognition particle consists of a 7S RNA molecule
CC of 300 nucleotides and six protein subunits: SRP72, SRP68, SRP54,
CC SRP19, SRP14 and SRP9.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- SIMILARITY: Belongs to the SRP9 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).
CC
CC EMBL; X78304; CAA55114.1; -;
CC EMBL; AK011720; BAB27800.1; -;
CC EMBL; AK020583; BAB32138.1; -;
CC EMBL; AK020620; BAB32151.1; -;
CC EMBL; BC039648; AAB39648.1; -;
CC PIR; S57500; S57500.
CC PDB; 1914; X-ray; @=-;
CC MGD; MG1:1350930; SRP9.
CC InterPro; IPR008832; SRP9.
CC InterPro; IPR009018; SRP9/14.
CC Pfam; PF05486; SRP9; 1.
CC 3D-structure; RNA-binding; Signal recognition particle.
CC INIT MET 0 0 By similarity
CC CONFLICT 21 21 P -> S (in Ref. 2; BAB27800).
CC HELIX 6 19
CC TURN 21 23
CC STRAND 25 31
CC FT

| | | | |
|----|----------|------------------|-------------------------|
| FT | TURN | 32 | 35 |
| FT | STRAND | 36 | 42 |
| FT | STRAND | 47 | 52 |
| FT | STRAND | 54 | 54 |
| FT | HELIX | 55 | 57 |
| FT | HELIX | 58 | 73 |
| SQ | SEQUENCE | 85 AA; 10063 MW; | 01C529D4714C62AD CRC64; |

Query Match 81.6%; Score 62; DB 1; Length 85;
Best Local Similarity 100.0%; Pred.No. 0.0063;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | |
|----|-----------------|--------|
| QY | 3 EEF5RAAEKLYLA | 15 |
| DB | 7 EEF5RAAEKLYLA | 19 |

```

RESULT 9
AAH66957 PRELIMINARY; PRT; 86 AA.
AC AAH66957;
DT 03-MAR-2004 (TEMBLrel. 27, Created)
DT 03-MAR-2004 (TEMBLrel. 27, Last sequence update)
DT 03-MAR-2004 (TEMBLrel. 27, Last annotation update)
DE Signal recognition particle 9kDa.
GN SRP9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxId=9606;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Brain;
RC MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.J., Jordan H., Moore T., Max S.I., Wang J., Heise F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield V.S.,
RA Krzywicki M.I., Skalka U., Smalusz D.E., Schnerch A., Schein J.E.,
RA Jones S.U., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RA RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RN SEQUENCE FROM N.A.
RP TISSUE=Brain;
RC Strausberg R.;
RA Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC066957; AAH66957.1; -.
SQ SEQUENCE 86 AA; 10112 MW; D82CA071AF0B6F65 CRC64;
```

Query Match 81.6%; Score 62; DB 2; Length 86;
Best Local Similarity 100.0%; Pred.No. 0.0063;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | |
|----|-----------------|--------|
| QY | 3 EEF5RAAEKLYLA | 15 |
| DB | 8 EEF5RAAEKLYLA | 20 |

```

RESULT 10
BAC34557
```

ID BAC34557 PRELIMINARY; PRT; 86 AA.
AC BAC34557;
DT 14-APR-2004 (TrEMBLrel. 27, Created)
DT 14-APR-2004 (TrEMBLrel. 27, Last sequence update)
DE 12 days embryo spinal ganglion cDNA, RIKEN full-length enriched library, clone: D130016712 product: signal recognition particle 9 kDa, full insert sequence.
DE full insert sequence.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
RL Meth. Enzymol. 303:19-44(1999).
RN [4]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630(2000).
RN [5]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P., Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M., Itoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer."
RL Genome Res. 10:1757-1771(2000).
RN [6]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P., Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirazane T., Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T., Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y., Saito K., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T., Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK051208; BAC34557.1; -.
SQ SEQUENCE 86 AA; 10194 MW; 655860497132AD2C CRC64;
Query Match 81.6%; Score 62; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.0063;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 EEFSSRAAEKLYIA 15
| | | | | | | | | | | | | | | | | | | | | |
Db 8 EEFSSRAAEKLYIA 20
| | | | | | | | | | | | | | | | | | | | | |
RESULT 11
Q7ZW94
ID Q7ZW94 PRELIMINARY; PRT; 86 AA.
AC Q7ZW94;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE 01-VAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to signal recognition particle 9kDa.
GN Name=sgc:56668;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Whole body;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Buetow K.H., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullah S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyx S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smalusz D.E., Schnerch A., Schein J.E., Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RN SEQUENCE FROM N.A.
RC TISSUE=Whole body;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC049522; AAH49522.1; -.
DR HSSP; P49458; I280.
DR GO; GO:0005786; C:signal recognition particle; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0045900; P:negative regulation of translational elonga...; IEA.
DR GO; GO:0006605; P:protein targeting; IEA.
DR InterPro; IPR008832; SRP.
DR InterPro; IPR009018; SRP9/14.
DR Pfam; PF05486; SRP9; 1.
SQ SEQUENCE 86 AA; 10076 MW; 65853A3ABA7FAF33 CRC64;
Query Match 72.4%; Score 55; DB 2; Length 86;
Best Local Similarity 91.7%; Pred. No. 0.11;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 3 EEFSSRAAEKLYL 14
| | | | | | | | | | | | | | | | | | | | | |
Db 8 EEFSSRAAEKLYL 19
| | | | | | | | | | | | | | | | | | | | | |

| RESULT 12 | Q81Y91 | PRELIMINARY; | PRT; | 464 AA. |
|--------------------------|--|---|---------|---------|
| DT | Q81Y91 | 30-MAY-2000 (Rel. 39, Created) | | |
| DT | Q81Y91 | 30-MAY-2000 (Rel. 39, Last sequence update) | | |
| DT | Q81Y91 | 05-JUL-2004 (Rel. 44, Last annotation update) | | |
| DE | Cyclin A1. | | | |
| GN | Name=CCNA1; | | | |
| OS | Homo sapiens (Human) | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | |
| OX | NCBI_TaxID=9606; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | TISSUE=Brain; | | | |
| RX | MEDLINE=22388257; PubMed=12477932; | | | |
| RA | Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., | | | |
| RA | Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., | | | |
| RA | Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., | | | |
| RA | Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., | | | |
| RA | Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong J., | | | |
| RA | Crowton M., Soares M.B., Bonaldi M.P., Casavant T.L., Scheetz T.E., | | | |
| RA | Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C., | | | |
| RA | Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., | | | |
| RA | Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., | | | |
| RA | Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., | | | |
| RA | Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., | | | |
| RA | Fahy J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A., | | | |
| RA | Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., | | | |
| RA | Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., | | | |
| RA | Rodríguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., | | | |
| RA | Krzywinski M.I., Skalska U., Smitt J.E., Schnerch A., Schein J.E., | | | |
| RA | Jones S.J., Marra M.A., | | | |
| RT | "Generation and initial analysis of more than 15,000 full-length human | | | |
| RT | and mouse cDNA sequences." | | | |
| RL | Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | TISSUE=Brain; | | | |
| RA | Strausberg R.; | | | |
| RL | Submitted (AUG-2002) to the EMBL/GenBank/DBSJ databases. | | | |
| CC | -1- SIMILARITY: Belongs to the cyclin family. | | | |
| CC | EMBL; BC036346; ABH36346.1; -- | | | |
| DR | HSSP; P20248; 1FIN. | | | |
| DR | GO; GO:0005634; C:nucleus; IEA. | | | |
| DR | GO; GO:000074; P:regulation of cell cycle; IEA. | | | |
| DR | InterPro; IPR006670; Cyclin. | | | |
| DR | InterPro; IPR004367; Cyclin Cterm. | | | |
| DR | InterPro; IPR011028; Cyclin-like. | | | |
| DR | InterPro; IPR006671; Cyclin_N. | | | |
| DR | Pfam; PF02984; Cyclin_C; 1. | | | |
| DR | Pfam; PF00334; Cyclin_N; 1. | | | |
| DR | SMART; SM00385; Cyclin; 2. | | | |
| DR | PROSITE; PS00292; Cyclin; 1. | | | |
| KW | Cell cycle; Cell division; Cyclin. | | | |
| SQ | SEQUENCE 464 AA; 52229 MW; E67C656D741B7D73 CRC64; | | | |
| Query Match | 65.8%; Score 50; DB 2; Length 464; | | | |
| Best Local Similarity | 62.5%; Pred. No. 4.8; | | | |
| Matches 10; Conservative | 2; Mismatches 4; Indels 0; Gaps 0; | | | |
| Qy | 1 IGEEFSRAAEKLYLAV 16 | | | |
| Db | 253 VGEYKLRATLYLAV 268 | | | |
| RESULT 13 | | | | |
| CGAL_HUMAN | STANDARD; | PRT; | 465 AA. | |
| CGAL_HUMAN | CGAL_HUMAN | | | |

Query Match 65.8%; Score 50; DB 1; Length 465;
 Best Local Similarity 62.5%; Pred. No. 4.8;
 Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 IGEEFSPRAAEKLYLAV 16
 :|||: |||||
 Db 254 VGEYKLRATLYLAV 269

RESULT 14

QY73L6 PRELIMINARY; PRT; 390 AA.
 AC QY73L6;
 DT 01-OCT-2003 (T-EMBLrel. 25, Created)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
 DE Cyclin A1.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bauer M.P., Goetz F.W.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to the cyclin family.
 DR EMBL: AF268045; AAP47015.1; -;
 DR GO: GO:0005634; C:nucleus; IEA.
 DR GO: GO:0000074; P:regulation of cell cycle; IEA.
 DR InterPro; IPR006670; Cyclin.
 DR InterPro; IPR004367; Cyclin_Cterm.
 DR InterPro; IPR006671; Cyclin_N.
 DR Pfam; PF02984; Cyclin_C; 1.
 DR Pfam; PF00134; Cyclin_N; 1.
 DR SMART; SM00385; CYCLIN; 2.
 DR PROSITE; PS00292; CYCLINS; 1.
 KW Cell cycle; Cell division; Cyclin.
 SQ SEQUENCE 390 AA; 43966 MW; 79DD1F8F1201FF9D CRC64;

Query Match 63.2%; Score 48; DB 2; Length 390;
 Best Local Similarity 56.2%; Pred. No. 9.1;
 Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 IGEEFSPRAAEKLYLAV 16
 :|||: :|||
 Db 179 VGEYKLCSETLYLAV 194

RESULT 15

QY83P3 PRELIMINARY; PRT; 395 AA.
 AC QY83P3;
 DT 01-MAR-2003 (T-EMBLrel. 23, Created)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
 DE Solute binding protein of ABC transporter for branched-chain amino acids.
 GN OrderedLocustNames=BL1714;
 OS Bifidobacterium longum.
 OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
 OC Bifidobacteriaceae; Bifidobacterium.
 OX NCBI_TaxID=216816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=NC 2705;
 RX MEDLINE=22294977; PubMed=12381787;
 RA Schell M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,
 RA Pessi G., Zwaalen M.-C., Desiere F., Bork P., Delley M.,
 RA Pidmore R.D., Arigoni F.;
 RT "The genome sequence of Bifidobacterium longum reflects its adaptation to the human gastrointestinal tract."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).

DR EMBL: AE014805; AAN25498.1; -;
 DR GO: GO:0030288; C:periplasmic space (sensu Gram-negative bact...; IEA.
 DR GO: GO:0005279; P:amino acid-polyamine transporter activity; IEA.
 DR GO: GO:0006865; P:amino acid transport; IEA.
 DR InterPro; IPR001828; ANF_receptor.
 DR InterPro; IPR000709; Leu_ile_val_bind.
 DR Pfam; PF01094; ANF_receptor; 1.
 DR PRINTS; PR00337; LEUILEVALBP.
 KW Complete proteome.
 SQ SEQUENCE 395 AA; 41399 MW; BFB76ABDF722D6A8 CRC64;

Query Match 63.2%; Score 48; DB 2; Length 395;
 Best Local Similarity 62.5%; Pred. No. 9.2;
 Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 IGEEFSPRAAEKLYLAV 16
 :|||: :|||
 Db 196 VGEAFAPAAEKLGLEV 211

Search completed: November 10, 2004, 15:53:39
 Job time : 28.8954 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:55:47 ; Search time 10.6406 Seconds
(without alignments)
143.349 Million cell updates/sec

Title: US-10-092-750-226

Perfect score: 117
Sequence: 1 KAEVQIARKLQCIADQFHLHLVL 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```
1: /cgn2_6/prodata/1/iaa/5A_COMB.pcp.*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pcp.*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pcp.*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pcp.*
5: /cgn2_6/prodata/1/iaa/PCrTru_COMB.pcp.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pcp.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query | | | DB | ID | Description |
|------------|-------|-------|--------|----|----------------------|-------------------|
| | Score | Match | Length | | | |
| 1 | 48 | 41.0 | 286 | 4 | US-09-270-767-57322 | Sequence 57322, A |
| 2 | 48 | 41.0 | 764 | 4 | US-09-270-767-42064 | Sequence 42064, A |
| 3 | 44.5 | 38.0 | 229 | 4 | US-09-252-991A-30629 | Sequence 30629, A |
| 4 | 44 | 37.6 | 806 | 4 | US-09-270-767-34213 | Sequence 34213, A |
| 5 | 44 | 37.6 | 806 | 4 | US-09-270-767-49430 | Sequence 49430, A |
| 6 | 43.5 | 37.2 | 507 | 4 | US-09-543-681A-5913 | Sequence 5913, Ap |
| 7 | 43.5 | 37.2 | 947 | 3 | US-09-228-986-73 | Sequence 73, Appl |
| 8 | 43.5 | 37.2 | 947 | 4 | US-10-101-464A-73 | Sequence 73, Appl |
| 9 | 43.5 | 37.2 | 1621 | 1 | US-08-242-677-2 | Sequence 2, Appl |
| 10 | 43 | 36.8 | 384 | 4 | US-09-533-029-50 | Sequence 50, Appl |
| 11 | 43 | 36.8 | 440 | 3 | US-09-100-664A-3 | Sequence 3, Appl |
| 12 | 43 | 36.8 | 440 | 4 | US-09-335-983-3 | Sequence 3, Appl |
| 13 | 43 | 36.8 | 440 | 4 | US-09-553-867A-3 | Sequence 3, Appl |
| 14 | 42 | 35.9 | 233 | 4 | US-09-270-767-41050 | Sequence 41050, A |
| 15 | 42 | 35.9 | 233 | 4 | US-09-270-767-56266 | Sequence 56266, A |
| 16 | 42 | 35.9 | 440 | 3 | US-09-100-664A-2 | Sequence 2, Appl |
| 17 | 42 | 35.9 | 440 | 3 | US-09-100-664A-4 | Sequence 4, Appl |
| 18 | 42 | 35.9 | 440 | 4 | US-09-335-983-2 | Sequence 2, Appl |
| 19 | 42 | 35.9 | 440 | 4 | US-09-335-983-4 | Sequence 4, Appl |
| 20 | 42 | 35.9 | 440 | 4 | US-09-553-867A-2 | Sequence 2, Appl |
| 21 | 42 | 35.9 | 440 | 4 | US-09-553-867A-4 | Sequence 4, Appl |
| 22 | 41 | 35.0 | 99 | 3 | US-09-199-637A-75 | Sequence 75, Appl |
| 23 | 41 | 35.0 | 124 | 4 | US-09-513-999C-5335 | Sequence 5335, Ap |
| 24 | 41 | 35.0 | 277 | 4 | US-09-710-279-94 | Sequence 94, Appl |
| 25 | 41 | 35.0 | 316 | 3 | US-09-134-001C-3736 | Sequence 3736, Ap |
| 26 | 41 | 35.0 | 399 | 4 | US-09-813-818-4 | Sequence 4, Appl |
| 27 | 41 | 35.0 | 401 | 4 | US-09-813-818-3 | Sequence 3, Appl |

ALIGNMENTS

RESULT 1

```

US-09-270-767-57322
/ Sequence 57322, Application US/09270767
/ Patent No. 6703491
/ GENERAL INFORMATION:
/ APPLICANT: Homburger et al.
/ TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
/ FILE REFERENCE: File Reference: 7326-094
/ CURRENT APPLICATION NUMBER: US/09/270,767
/ CURRENT FILING DATE: 1999-03-17
/ NUMBER OF SEQ ID NOS: 62517
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 57322
/ LENGTH: 286
/ TYPE: PRT
/ ORGANISM: Drosophila melanogaster
/ FEATURE:
/ OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-57322

```

```

Query Match      41.0%; Score 48; DB 4; Length 286;
Best Local Similarity 50.0%; Pred. No. 4.7;
Matches 11; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

y      2 AEVQI--ARKLCQIADQFRLH 21
      |||: |||: |||: |||:
      192 AEIOVINRKVICISEOVIRLH 213

```

```

RESULT 2
US-09-270-767-42064
; Sequence 42064, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42064
; LENGTH: 764
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-42064

```

Query Match
41.0%; Score 48; DB 4; Length 764;

```

Best Local Similarity 50.0%; Pred. No. 14;
Matches 11; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

QY 2 AEVQI--ARKLQCIADQFRLH 21
   ||:|: || ||:|:|:|:|
Db 670 AEIQLNIRKYICISQVIRLH 691

RESULT 3
US-09-252-991A-30629
; Sequence 30629, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.135
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30629
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30629

Query Match 38.0%; Score 44.5; DB 4; Length 229;
Best Local Similarity 57.9%; Pred. No. 14;
Matches 11; Conservative 2; Mismatches 3; Indels 3; Gaps 1;

QY 1 KAEVQIARKL---QCIADQ 16
   ||:|:|:|:|:|
Db 155 KGEVRIARRLIGACQSPDQ 173

RESULT 4
US-09-270-767-34213
; Sequence 34213, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 34213
; LENGTH: 806
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-34213

Query Match 37.6%; Score 44; DB 4; Length 806;
Best Local Similarity 44.4%; Pred. No. 68;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 KAEVQIARKLQCIADQFH 18
   ||:|:|:|:|:|
Db 494 KIDIQSRKLKVLATYFH 511

RESULT 5
US-09-270-767-49430
; Sequence 49430, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:

```

```

; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 49430
; LENGTH: 806
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-49430

Query Match 37.6%; Score 44; DB 4; Length 806;
Best Local Similarity 44.4%; Pred. No. 68;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 KAEVQIARKLQCIADQFH 18
   ||:|:|:|:|:|
Db 494 KIDIQSRKLKVLATYFH 511

RESULT 6
US-09-543-681A-5913
; Sequence 5913, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5913
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5913

Query Match 37.2%; Score 43.5; DB 4; Length 507;
Best Local Similarity 40.0%; Pred. No. 49;
Matches 8; Conservative 6; Mismatches 3; Indels 3; Gaps 1;

QY 3 EVQIARKLQCIADQFRLHV 22
   ||:|:|:|:|:|
Db 240 EMVGRKLE---DQYPRHI 256

RESULT 7
US-09-228-986-73
; Sequence 73, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 73
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-228-986-73

```

```

Query Match          37.2%; Score 43.5; DB 3; Length 947;
Best Local Similarity 43.5%; Pred. No. 99;
Matches 10; Conservative 4; Mismatches 6; Indels 3; Gaps 1;

QY      3 EVQIARKLQ---CIADQFHLHV 22
      |||:||||:| |||:
Db      636 EVQLARKIEGFIHPDSIHLRI 658

RESULT 8
US-10-101-464A-73
; Sequence 73, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-73

Query Match          37.2%; Score 43.5; DB 4; Length 947;
Best Local Similarity 43.5%; Pred. No. 99;
Matches 10; Conservative 4; Mismatches 6; Indels 3; Gaps 1;

QY      3 EVQIARKLQ---CIADQFHLHV 22
      |||:||||:| |||:
Db      636 EVQLARKIEGFIHPDSIHLRI 658

RESULT 9
US-08-242-677-2
; Sequence 2, Application US/08242677
; Patent No. 5677143
; GENERAL INFORMATION:
; APPLICANT: Gaynor, Richard B
; APPLICANT: Wu, Foon W.
; TITLE OF INVENTION: Cellular Nucleic Acid Binding Protein
; TITLE OF INVENTION: and Uses Thereof in regulating Gene Expression and in the
; TITLE OF INVENTION: Treatment of AIDS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMEUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/242,677
; FILING DATE:

```

```

RESULT 14
US-09-270-767-41050
; Sequence 41050, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 41050
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-41050

Query Match          35.9%; Score 42; DB 4; Length 233;
Best Local Similarity 36.0%; Pred. No. 36;
Matches 9; Conservative 5; Mismatches 5; Indels 6; Gaps 1;

Qy      3  EVOIARKLQ-----CIADQFRLH 21
      : : : :
      : : : :
Db      100  KVNLIINLRLOKCNVXCIYEQYRLH 124
      : : : :

```



```

RESULT 15
US-09-270-767-56266
; Sequence 56266, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56266
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-56266

Query Match      35.9%; Score 42; DB 4; Length 233;
Best Local Similarity 36.0%; Pred No. 36;
Matches 9; Conservative 5; Mismatches 5; Indels 6; Gaps 1;

QY      3 EVOIARKLQ-----CIADQFHLH 21
      :|: :|||
Db      100 KVNLIINRLQKCNVWVCIVEQYRRLH 124

Search completed: November 10, 2004, 15:57:24
Job time : 11.6406 secs

```

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 02:43:24 ; Search time 55.5765 Seconds
(without alignments)
146.426 Million cell updates/sec

Title: US-10-092-750-226

Perfect score: 117

Sequence: 1 KAEVQIARKLQCIADQFRLHVL 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1568699 seqs, 33819137 residues

Total number of hits satisfying chosen parameters: 1568699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 117 | 100.0 | 23 | 14 | US-10-092-750-226 |
| 2 | 57 | 48.7 | 157 | 17 | US-10-739-930-10801 |
| 3 | 53 | 45.3 | 79 | 17 | US-10-425-115-329239 |
| 4 | 53 | 45.3 | 250 | 17 | US-10-425-115-244863 |
| 5 | 52 | 44.4 | 97 | 16 | US-10-767-701-60849 |
| 6 | 52 | 44.4 | 113 | 17 | US-10-425-115-217289 |
| 7 | 51 | 43.6 | 89 | 16 | US-10-767-701-42415 |
| 8 | 47.5 | 40.6 | 65 | 17 | US-10-425-115-276339 |
| 9 | 47 | 40.2 | 437 | 15 | US-10-424-599-233142 |
| 10 | 46 | 39.3 | 519 | 16 | US-10-437-963-157663 |
| 11 | 45.5 | 38.9 | 119 | 11 | US-09-864-408A-5252 |
| 12 | 45 | 38.5 | 62 | 17 | US-10-425-115-218585 |
| 13 | 45 | 38.5 | 205 | 15 | US-10-424-599-202266 |

| | | | | | | |
|----|------|------|------|----|----------------------|--------------------|
| 14 | 45 | 38.5 | 355 | 15 | US-10-424-599-197587 | Sequence 197587, A |
| 15 | 45 | 38.5 | 358 | 15 | US-10-425-114-38292 | Sequence 38292, A |
| 16 | 45 | 38.5 | 460 | 14 | US-10-278-173-100 | Sequence 100, App |
| 17 | 45 | 38.5 | 460 | 14 | US-10-278-536-84 | Sequence 84, Appl |
| 18 | 45 | 38.5 | 460 | 14 | US-10-225-066A-608 | Sequence 608, App |
| 19 | 45 | 38.5 | 460 | 14 | US-10-302-267-25 | Sequence 26, Appl |
| 20 | 45 | 38.5 | 460 | 15 | US-10-374-780A-2362 | Sequence 2362, Ap |
| 21 | 45 | 38.5 | 460 | 15 | US-10-413-699B-376 | Sequence 376, App |
| 22 | 45 | 38.5 | 475 | 16 | US-10-437-963-181315 | Sequence 181315, A |
| 23 | 45 | 38.5 | 869 | 16 | US-10-437-963-174442 | Sequence 174442, A |
| 24 | 44.5 | 38.0 | 60 | 17 | US-10-425-115-263998 | Sequence 263998, A |
| 25 | 44 | 37.6 | 55 | 16 | US-10-437-963-132580 | Sequence 132580, A |
| 26 | 44 | 37.6 | 157 | 15 | US-10-424-599-196483 | Sequence 196483, A |
| 27 | 44 | 37.6 | 385 | 15 | US-10-425-114-41842 | Sequence 41842, A |
| 28 | 44 | 37.6 | 392 | 17 | US-10-739-930-8800 | Sequence 8800, Ap |
| 29 | 44 | 37.6 | 491 | 15 | US-10-424-599-196481 | Sequence 196481, A |
| 30 | 44 | 37.6 | 588 | 15 | US-10-424-599-163333 | Sequence 163333, A |
| 31 | 43.5 | 37.2 | 947 | 14 | US-10-101-464A-73 | Sequence 73, Appl |
| 32 | 43.5 | 37.2 | 1621 | 16 | US-10-408-765A-2038 | Sequence 2038, App |
| 33 | 43 | 36.8 | 105 | 14 | US-10-029-385-33323 | Sequence 33323, A |
| 34 | 43 | 36.8 | 121 | 17 | US-10-425-115-212000 | Sequence 212000, A |
| 35 | 43 | 36.8 | 130 | 16 | US-10-437-963-141803 | Sequence 141803, A |
| 36 | 43 | 36.8 | 172 | 15 | US-10-424-599-284812 | Sequence 284812, A |
| 37 | 43 | 36.8 | 172 | 15 | US-10-425-114-42330 | Sequence 42330, A |
| 38 | 43 | 36.8 | 172 | 15 | US-10-425-114-53361 | Sequence 53361, A |
| 39 | 43 | 36.8 | 192 | 16 | US-10-767-701-62916 | Sequence 62916, A |
| 40 | 43 | 36.8 | 214 | 16 | US-10-437-963-133369 | Sequence 133369, A |
| 41 | 43 | 36.8 | 285 | 17 | US-10-425-115-318633 | Sequence 318633, A |
| 42 | 43 | 36.8 | 283 | 15 | US-10-282-122A-67725 | Sequence 67725, A |
| 43 | 43 | 36.8 | 283 | 15 | US-10-282-122A-70004 | Sequence 70004, A |
| 44 | 43 | 36.8 | 313 | 15 | US-10-425-114-38329 | Sequence 38329, A |
| 45 | 43 | 36.8 | 321 | 15 | US-10-424-599-284810 | Sequence 284810, A |

ALIGNMENTS

RESULT 1
US-10-092-750-226
; Sequence 226 Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; PRIOR FILING DATE: 2002-03-07
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 226
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-226

Query Match 100.0%; Score 117; DB 14; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.2e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KAEVQIARKLQCIADQFRLHVL 23
DB 1 KAEVQIARKLQCIADQFRLHVL 23

RESULT 2
US-10-739-930-10801
; Sequence 10801, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
 ; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
 ; FILE REFERENCE: 38-21(53377)B
 ; CURRENT APPLICATION NUMBER: US/10/739,930
 ; CURRENT FILING DATE: 2003-12-18
 ; NUMBER OF SEQ ID NOS: 11088
 ; SEQ ID NO 10801
 ; LENGTH: 157
 ; TYPE: PRT
 ; ORGANISM: Triticum aestivum
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: TRIAE-23APR03-C71899_1.p
 US-10-739-930-10801

Query Match 48.7%; Score 57; DB 17; Length 157;
 Best Local Similarity 57.9%; Pred. No. 0.53; Mismatches 3; Indels 5; Gaps 0;
 Matches 11; Conservative 3; Mismatches 0; Gaps 0;

QY 5 QIARKLQCIADQPHRLHVL 23
 DB 116 QMARSLAVSDYFHLRL 134

RESULT 3
 US-10-425-115-329239
 ; Sequence 329239, Application US/10425115
 ; Publication No. US20040214272A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: PLANTS
 ; FILE REFERENCE: 38-21(53222)B
 ; CURRENT APPLICATION NUMBER: US/10/425,115
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 369326
 ; SEQ ID NO 329239
 ; LENGTH: 79
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: MRT4577_63339C.1.p
 US-10-425-115-329239

Query Match 45.3%; Score 53; DB 17; Length 79;
 Best Local Similarity 63.2%; Pred. No. 1.1;
 Matches 12; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 QIARKLQCIADQPHRLHVL 23
 DB 41 QAARGLLAIADYFHLRLAL 59

RESULT 4
 US-10-425-115-244863
 ; Sequence 244863, Application US/10425115
 ; Publication No. US20040214272A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: PLANTS
 ; FILE REFERENCE: 38-21(53222)B
 ; CURRENT APPLICATION NUMBER: US/10/425,115
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 369326
 ; SEQ ID NO 244863
 ; LENGTH: 250

; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: MRT4577_154904C.1.p
 US-10-425-115-244863

Query Match 45.3%; Score 53; DB 17; Length 250;
 Best Local Similarity 52.6%; Pred. No. 3.7;
 Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 5 QIARKLQCIADQPHRLHVL 23
 DB 218 QMARSLAVSDYFHLRL 236

RESULT 5
 US-10-767-701-60849
 ; Sequence 60849, Application US/10767701
 ; Publication No. US20040172684A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
 ; FILE REFERENCE: 38-21(53535)B
 ; CURRENT APPLICATION NUMBER: US/10/767,701
 ; CURRENT FILING DATE: 2004-01-29
 ; NUMBER OF SEQ ID NOS: 63128
 ; SEQ ID NO 60849
 ; LENGTH: 97
 ; TYPE: PRT
 ; ORGANISM: Sorghum bicolor
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 9298346.p
 US-10-767-701-60849

Query Match 44.4%; Score 52; DB 16; Length 97;
 Best Local Similarity 63.2%; Pred. No. 2;
 Matches 12; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 QIARKLQCIADQPHRLHVL 23
 DB 62 QAARCLLAIADYFHLRLAL 80

RESULT 6
 US-10-425-115-217289
 ; Sequence 217289, Application US/10425115
 ; Publication No. US20040214272A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: PLANTS
 ; FILE REFERENCE: 38-21(53222)B
 ; CURRENT APPLICATION NUMBER: US/10/425,115
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 369326
 ; SEQ ID NO 217289
 ; LENGTH: 113
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: MRT4577_129769C.1.p
 US-10-425-115-217289

Query Match 44.4%; Score 52; DB 17; Length 113;
 Best Local Similarity 45.0%; Pred. No. 2.3;
 Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

RESULT 9
US-10-424-598-233142
; Sequence 233142, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

RESULT 11
US-09-864-408A-5252
; Sequence 5252, Application US/09864408A
; Publication No. US2004009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.

Best Local Similarity 52.6%; Pred. No. 1e+02;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 5 QIARKLQCIADQGFHRIHVL 23
| | | | |
| | | | |
| | | | |
Db 328 QAARCLVAISEYFHRRLAL 346

Search completed: November 11, 2004, 07:41:47
Job time : 55.5765 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:52:32 ; Search time 7.85765 Seconds
(without alignments)
281.634 Million cell updates/sec

Title: US-10-092-750-226

Perfect score: 117
Sequence: 1 KAEVQIARKLQCIADQPHRLHL 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB | ID | Description |
|------------|-------|---------------|--------|----|--------|--------------------|
| 1 | 48 | 41.0 | 650 | 2 | T00094 | endostyle-specific |
| 2 | 45 | 38.5 | 218 | 2 | AD0112 | ribose-5-phosphate |
| 3 | 45 | 38.5 | 678 | 2 | G84682 | hypothetical prote |
| 4 | 45 | 38.5 | 688 | 2 | B84631 | hypothetical prote |
| 5 | 45 | 38.5 | 733 | 2 | E85357 | hypothetical prote |
| 6 | 45 | 38.5 | 747 | 2 | T52572 | cyclic nucleotide |
| 7 | 44.5 | 38.0 | 183 | 2 | D83069 | conserved hypothet |
| 8 | 44 | 37.6 | 2632 | 2 | T18718 | dynein heavy chain |
| 9 | 43.5 | 37.2 | 170 | 1 | CYRTG2 | gamma-crystallin 2 |
| 10 | 43.5 | 37.2 | 174 | 2 | C24060 | gamma-crystallin 2 |
| 11 | 43.5 | 37.2 | 174 | 2 | I83432 | gamma-C-crystallin |
| 12 | 43.5 | 37.2 | 368 | 2 | A97189 | lps biosynthesis p |
| 13 | 43.5 | 37.2 | 1621 | 2 | S62356 | TRP-185 protein - |
| 14 | 43 | 36.8 | 365 | 2 | JC7527 | nuclear retroviral |
| 15 | 43 | 36.8 | 384 | 2 | P86353 | protein P2E2.14 [i |
| 16 | 43 | 36.8 | 384 | 2 | S48523 | transcription fact |
| 17 | 43 | 36.8 | 449 | 2 | G98992 | hypothetical prote |
| 18 | 43 | 36.8 | 482 | 2 | T07014 | phosphoglycerate k |
| 19 | 43 | 36.8 | 738 | 2 | E86294 | hypothetical prote |
| 20 | 43 | 36.8 | 746 | 2 | H86330 | probable cyclic nu |
| 21 | 43 | 36.8 | 1101 | 2 | T33153 | hypothetical prote |
| 22 | 42.5 | 36.3 | 406 | 2 | T31778 | hypothetical prote |
| 23 | 42.5 | 36.3 | 212 | 2 | T05113 | hypothetical prote |
| 24 | 42 | 35.9 | 217 | 2 | T26669 | hypothetical prote |
| 25 | 42 | 35.9 | 228 | 2 | T47847 | hypothetical prote |
| 26 | 42 | 35.9 | 322 | 2 | T03802 | cyclic nucleotide |
| 27 | 42 | 35.9 | 323 | 2 | AY9936 | hypothetical prote |
| 28 | 42 | 35.9 | 367 | 2 | F82098 | riboflavin-specifi |
| 29 | 42 | 35.9 | 452 | 2 | T39566 | hypothetical prote |

hypothetical prote
glutelin II precurs
glutelin 2 precurs
glutelin type I pr
glutelin II precur
glutelin precursor
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
SOX-LZ - rainbow t
hypothetical prote
major capsid prote
probable spectrin
hypothetical prote
hypothetical prote

ALIGNMENTS

RESULT 1

T00094
endostyle-specific protein - sea squirt (Ciona intestinalis)
C/Species: Ciona intestinalis
C/Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 09-Jul-2004
C/Accession: T00094
R/Ogasawara, M.; Satoh, N.
submitted to the EMBL Data Library, February 1998
A/Description: Isolation and characterization of endostyle-specific genes in the Ascidia
A/Reference number: Z14109
A/Accession: T00094
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-650 <OGA>
A/Cross-references: UNIPROT:O44236; EMBL:AB010895
A/Experimental source: endostyle

Query Match 41.0% ; Score 48; DB 2; Length 650;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 KAEVQIARKLQCIADQPHRL 20

Db 189 KCEEDVADKLKCIADYHTKL 208

RESULT 2

AD0112
ribose-5-phosphate isomerase (EC 5.3.1.6) [imported] - Yersinia pestis (strain CO92)
C/Species: Yersinia pestis
C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C/Accession: AD0112
R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Irl, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, I. Nature 413, 523-527, 2001
A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A/Reference number: AB00001; MUID:21470413; PMID:11586360
A/Accession: AD0112
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-218 <KUR>
A/Cross-references: UNIPROT:Q8ZH8; GB:AL590842; PIDN:CAC89759.1; PID:gl5978986; GSPDB:GN

A/Gene: rpiA
C/Superfamily: Haemophilus influenzae ribose-5-phosphate isomerase
C/Keywords: intramolecular oxidoreductase; isomerase

Query Match 38.5% ; Score 45; DB 2; Length 218;
Best Local Similarity 55.6%; Pred. No. 13;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

Qy      6 IARKLQCIADQFHRLH 23
      ||||| :||| :|||
Db      109 IARKFCIADSKQGVYL 126

RESULT 3
hypothetical protein At2g28260 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: G84682
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: G84682
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-678 <STO>
A:Cross-references: UNIPROT:Q9SL29; GB:AE002093; NID:g4803955; PIDN:AAD29827.1; GSPDB:GN
C:Genetics:
A:Gene: At2g28260
A:Map position: 2

Query Match      38.5%; Score 45; DB 2; Length 678;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy      7 ARKLQCIADQFHRLH 21
      ||||| :||| :|||
Db      577 AEDLQFVASQFRLH 591

RESULT 4
hypothetical protein At2g23980 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: B84631
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: B84631
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-688 <STO>
A:Cross-references: UNIPROT:O82226; GB:AE002093; NID:g3738325; PIDN:AAC63666.1; GSPDB:GN
C:Genetics:
A:Gene: At2g23980
A:Map position: 2

Query Match      38.5%; Score 45; DB 2; Length 688;
Best Local Similarity 56.2%; Pred. No. 41;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy      6 IARKLQCIADQFHRLH 21
      ||||| :||| :|||
Db      560 IADELKFVASQFRLH 575

RESULT 5
hypothetical protein At4g30560 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: B85357
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999

A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: B85357
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-733 <STO>
A:Cross-references: UNIPROT:Q9MCA4; GB:NC_001268; NID:g7269957; PIDN:CAB9774.1; GSPDB:GN
C:Genetics:
A:Gene: AT4g30560
A:Map position: 4

Query Match      38.5%; Score 45; DB 2; Length 733;
Best Local Similarity 56.2%; Pred. No. 44;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy      6 IARKLQCIADQFHRLH 21
      ||||| :||| :|||
Db      618 IADELKFVASQFRLH 633

RESULT 6
cyclic nucleotide and calmodulin-regulated ion channel [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C>Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 09-Jul-2004
C:Accession: T52572
R:Kohler, C.; Merkle, T.; Neuhaus, G.
Plant J. 18, 97-104, 1999
A:Title: Characterisation of a novel gene family of putative cyclic nucleotide- and calmo
A:Reference number: T226120
A:Accession: T52572
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-747 <KO>
A:Cross-references: UNIPROT:O82226; EMBL:Y17914; PIDN:CAB40131.1
C:Genetics:
A:Gene: cngc6

Query Match      38.5%; Score 45; DB 2; Length 747;
Best Local Similarity 56.2%; Pred. No. 44;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy      6 IARKLQCIADQFHRLH 21
      ||||| :||| :|||
Db      619 IADELKFVASQFRLH 634

RESULT 7
conserved hypothetical protein PA4612 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: D83069
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bri
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lam,
Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: D83069
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-183 <STO>
A:Cross-references: UNIPROT:Q9HVH8; GB:AE004875; NID:g9950857; PIDN:AAG08000
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4612

Query Match      38.0%; Score 44.5; DB 2; Length 183;
Best Local Similarity 57.9%; Pred. No. 13;
Matches 11; Conservative 2; Mismatches 3; Indels 3; Gaps 1;

Qy      1 KAEVQIARKL---QCIADQ 16

```

```
Db      109 KGEVRIARLIGACQSPDQ 127
||:||||:||||
RESULT 8
T:18718
dynein heavy chain homolog - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18718; T23294
R: Sims, M.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19010
A:Accession: T18718
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2632 <WIL>
A:Cross-references: UNIPROT:P90736; EMBL:Z81028; PIDN:Z81028; GSPDB:GN00023; CESP:BO
A:Experimental source: clone B0365
R:Percy, C.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19722
A:Accession: T23294
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2632 <W12>
A:Cross-references: EMBL:Z81096; PIDN:CAB03163.1; GSPDB:GN00023; CESP:B0365.7
A:Experimental source: clone K03H4
C:Genetics:
A:Gene: CESP:B0365.7
A:Map position: 5
A:Introns: 75/2; 145/2; 312/1; 588/3; 636/1; 668/3; 744/3; 872/3; 1059/3; 1101/2; 1204/2
1; 2633/3
Query Match      37.6%; Score 44; DB 2; Length 2632;
Best Local Similarity 40.0%; Pred. No. 2.3e+02;
Matches 8; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY      1 KAEVQIARKLQCIADQFHRL 20
||:||||:||||:
Db      2316 KALLKYTEKLGSLDQFSQM 2335
||:||||:||||:
RESULT 9
CYRTG2
gamma-crystallin 2 - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change 09-Jul-2004
C:Accession: A02934
R: Moormann, R.J.M.; den Dunnen, J.T.; Bloemendal, H.; Schoenmakers, J.G.G.
Proc. Natl. Acad. Sci. U.S.A. 79, 6876-6880, 1982
A:Title: Extensive intragenic sequence homology in two distinct rat lens gamma-crystallin
A:Reference number: A93934; MUID:83091061; PMID:6294661
A:Accession: A02934
A:Molecule type: mRNA
A:Residues: 1-170 <MOO>
A:Cross-references: UNIPROT:P02529; GB:J00717; NID:G203632; PIDN:AAA40986.1; PID:G203633
C:Superfamily: beta-crystallin
C:Keywords: duplication; eye lens
F:1-36/Domain: crystallin repeat (fragment) <GK1>
F:37-79/Domain: crystallin repeat <GK2>
F:84-124/Domain: crystallin repeat <GK3>
F:125-164/Domain: crystallin repeat <GK4>
Query Match      37.2%; Score 43.5; DB 1; Length 170;
Best Local Similarity 36.0%; Pred. No. 17;
Matches 9; Conservative 5; Mismatches 6; Indels 5; Gaps 1;

QY      4 VOIARKLQCIADQFH-----RLHVL 23
||:||||:||||:
Db      99 MELSEDCSCIQDRFHLSEVRSLSHVL 123
||:||||:||||:
RESULT 10
C24060
gamma-crystallin 2-1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 21-Nov-1997
C:Accession: C24060
R: den Dunnen, J.T.; Moormann, R.J.M.; Lubsen, N.H.; Schoenmakers, J.G.G.
J. Mol. Biol. 189, 37-46, 1986
A:Title: Concerted and divergent evolution within the rat gamma-crystallin gene family.
A:Reference number: A92927; MUID:87060933; PMID:3783678
A:Accession: C24060
A:Molecule type: DNA
A:Residues: 1-174 <DEN>
A:Cross-references: GB:M19359; GB:M19354; GB:M19360; GB:M19361; GB:M19362; GB:M19363
C:Genetics:
A:Introns: 3/3; 84/3
C:Superfamily: beta-crystallin
C:Keywords: duplication
Query Match      37.2%; Score 43.5; DB 2; Length 174;
Best Local Similarity 36.0%; Pred. No. 18;
Matches 9; Conservative 5; Mismatches 6; Indels 5; Gaps 1;

QY      4 VOIARKLQCIADQFH-----RLHVL 23
||:||||:||||:
Db      103 MELSEDCSCIQDRFHLSEVRSLSHVL 127
||:||||:||||:
RESULT 11
I83432
gamma-C-crystallin - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 09-Jul-2004
C:Accession: I83432
R: den Dunnen, J.T.; van Neck, J.W.; Cremers, F.P.M.; Lubsen, N.H.; Schoenmakers, J.G.G.
Gene 78, 201-213, 1989
A:Title: Nucleotide sequence of the rat gamma-crystallin gene region and comparison with
A:Reference number: I60312; MUID:89378747; PMID:2777080
A:Accession: I83432
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-174 <RES>
A:Cross-references: UNIPROT:P02529; GB:M19359; NID:G203626; PIDN:AAA40983.1; PID:G203629
C:Genetics:
A:Gene: CRY-gamma-C
A:Introns: 3/3; 84/3
C:Superfamily: beta-crystallin
C:Keywords: duplication
Query Match      37.2%; Score 43.5; DB 2; Length 174;
Best Local Similarity 36.0%; Pred. No. 18;
Matches 9; Conservative 5; Mismatches 6; Indels 5; Gaps 1;

QY      4 VOIARKLQCIADQFH-----RLHVL 23
||:||||:||||:
Db      103 MELSEDCSCIQDRFHLSEVRSLSHVL 127
||:||||:||||:
RESULT 12
A97189
lps biosynthesis protein, RfB family [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C:Accession: A97189
R: Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clos
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: A97189
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-368 <KUR>
```

Db 25 QLYVGKALQLIRDQFPKLRII 45

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:50:40 ; Search time 38.5516 Seconds
(without alignments)
343.270 Million cell updates/sec

Title: US-10-092-750-226
Perfect score: 117
Sequence: 1 KAEVQIARKLQCIADQPHRLHVL 23

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB | ID | Description |
|------------|-------|---------------|--------|----|------------|--------------------|
| 1 | 109 | 93.2 | 184 | 2 | Q6NT30 | Q6nt30 homo sapien |
| 2 | 109 | 93.2 | 184 | 2 | Q6NT56 | Q6nt56 homo sapien |
| 3 | 109 | 93.2 | 184 | 2 | Q96LC9 | Q96lc9 homo sapien |
| 4 | 109 | 93.2 | 184 | 2 | AAH69328 | AAh69328 homo sapi |
| 5 | 109 | 93.2 | 184 | 2 | AAH69505 | AAh69505 homo sapi |
| 6 | 109 | 93.2 | 204 | 2 | Q9H7K7 | Q9h7k7 homo sapien |
| 7 | 109 | 93.2 | 261 | 2 | Q6P9F6 | Q6p9f6 homo sapien |
| 8 | 109 | 93.2 | 261 | 2 | AAH60783 | AAh60783 homo sapi |
| 9 | 109 | 93.2 | 261 | 2 | AAH70043 | AAh70043 homo sapi |
| 10 | 107 | 91.5 | 185 | 2 | Q8X589 | Q8x589 ratus narv |
| 11 | 106 | 90.6 | 185 | 2 | Q91ZE9 | Q91ze9 mus musculu |
| 12 | 106 | 90.6 | 271 | 2 | Q8BUK0 | Q8buk0 mus musculu |
| 13 | 51 | 43.6 | 269 | 2 | Q938D4 | Q938d4 mycobacteri |
| 14 | 49 | 41.9 | 325 | 2 | Q7U7D4 | Q7u7d4 leptococc |
| 15 | 48 | 41.0 | 351 | 2 | Q7Q279 | Q7q279 leptospira |
| 16 | 48 | 41.0 | 351 | 2 | Q8F5Y3 | Q8f5y3 leptospira |
| 17 | 48 | 41.0 | 351 | 2 | AAH70805 | AAh70805 leptospir |
| 18 | 48 | 41.0 | 650 | 2 | O44236 | O44236 ciona intes |
| 19 | 47 | 40.2 | 536 | 2 | Q8WSY4 | Q8wsy4 physarum po |
| 20 | 47 | 40.2 | 726 | 2 | Q8VVH8 | Q8vvh8 erwinia pyr |
| 21 | 47 | 40.2 | 1274 | 2 | Q45563 | Q45563 bacillus su |
| 22 | 46 | 39.3 | 219 | 1 | RPIA_NITEU | RpiA_NITEU |
| 23 | 46 | 39.3 | 261 | 2 | Q7QJZ6 | Q7qjz6 bacillus am |
| 24 | 46 | 39.3 | 261 | 2 | Q93158 | Q93158 bacillus su |
| 25 | 46 | 39.3 | 261 | 2 | CAE11252 | CAe11252 bacillus |
| 26 | 46 | 39.3 | 265 | 2 | Q6X892 | Q6x892 manis penta |
| 27 | 46 | 39.3 | 265 | 2 | AAQ57530 | AAq57530 manis pen |
| 28 | 46 | 39.3 | 317 | 2 | Q9NTV3 | Q9ntv3 homo sapien |
| 29 | 46 | 39.3 | 469 | 2 | Q8TC11 | Q8tc11 homo sapien |
| 30 | 46 | 39.3 | 617 | 2 | Q8RYH6 | Q8ryh6 oryza sativ |
| 31 | 46 | 39.3 | 1126 | 2 | Q6BTM6 | Q6btm6 debaryomyce |

32 45 38.5 101 2 Q9AG65 Q9ag65 spirochaeta
33 45 38.5 218 1 R1AL_YERPE R1al_yerpe
34 45 38.5 268 2 Q6X978 Q6x978 ochotona pr
35 45 38.5 268 2 AAQ57544 Aaq57544 ochotona pr
36 45 38.5 313 2 Q9FL60 Q9fl60 arabidopsis
37 45 38.5 365 2 P96033 P96033 synechococc
38 45 38.5 417 2 Q93XM5 Q93xm5 arabidopsis
39 45 38.5 441 2 Q89FP7 Q89fp7 bradyrhizob
40 45 38.5 467 2 Q6H434 Q6h434 oryza sativ
41 45 38.5 634 2 Q8PHV1 Q8phv1 xanthomonas
42 45 38.5 678 1 CG15_ARATH CG15_arath
43 45 38.5 733 1 CNG9_ARATH Cng9_arath
44 45 38.5 747 1 CNG26_ARATH Cng26_arath
45 45 38.5 806 2 Q76IN8 Q76in8 ciona intes

ALIGNMENTS

RESULT 1
Q6NT30 PRELIMINARY; PRT; 184 AA.
ID Q6NT30
AC Q6NT30;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE BC12 modifying factor.
GN Name=BMF;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PCR rescued clones;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.J., Wang J., Heien F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Boraldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Wolley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting M., Madan A., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Fahey J., Heiton E., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=PCR rescued clones;
RA Strausberg R.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC069505; AAH69505.1; -.
SQ SEQUENCE 184 AA; 20477 MW; 20178F7489662846 CRC64;
Query Match 93.2%; Score 109; DB 2; Length 184;
Best Local Similarity 95.5%; Pred. No. 4.6e-09;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KAEVQIARKLQCIADQPHRLHVL 22

Db 128 KAEVQIARKLQCIADQPHRLHVL 149

```

RESULT 2
Q6NT56 PRELIMINARY; PRT; 184 AA.
AC Q6NT56;
DT 03-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE BMF protein.
GN Name=BMF;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN
RP SEQUENCE FROM N.A.
RC TISSUE=PCR rescued clones;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Prange C.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=PCR rescued clones;
RA Strausberg R.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC069328; AAH69328.1; -.
SQ SEQUENCE 184 AA; 20480 MW; 21454B241C717853 CRC64;

Query Match 93.2%; Score 109; DB 2; Length 184;
Best Local Similarity 95.5%; Pred. No. 4.6e-09;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KAEVQIARKLQCIADQFHLHV 22
Db 128 QAEVQIARKLQCIADQFHLHV 149

RESULT 3
Q96LC9 PRELIMINARY; PRT; 184 AA.
AC Q96LC9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Bcl-2 modifying factor.
GN Name=BMF;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN
RP SEQUENCE FROM N.A.
RC TISSUE=PCR rescued clones;
RX MEDLINE=21432178; PubMed=11546872;
RA Puthalakath H., Villunger A., O'Reilly L.A., Beaumont J.G.,
RA Coultas L., Cheney R.E., Huang D.C., Strasser A.;
RT "Bmf: a proapoptotic Bcl-2-only protein regulated by interaction with

```

```

RT the myosin V actin motor complex, activated by anolakis."
RL Science 293:1829-1832(2001).
DR EMBL; AY029254; AAH38748.1; -.
SQ SEQUENCE 184 AA; 20508 MW; 21479B25CC727853 CRC64;

Query Match 93.2%; Score 109; DB 2; Length 184;
Best Local Similarity 95.5%; Pred. No. 4.6e-09;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KAEVQIARKLQCIADQFHLHV 22
Db 128 QAEVQIARKLQCIADQFHLHV 149

RESULT 4
AAH69328 PRELIMINARY; PRT; 184 AA.
AC AAH69328;
DT 20-MAY-2004 (TrEMBLrel. 27, Created)
DT 20-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DT 20-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE BMF protein.
GN BMF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN
RP SEQUENCE FROM N.A.
RC TISSUE=PCR rescued clones;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Prange C.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=PCR rescued clones;
RA Strausberg R.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC069328; AAH69328.1; -.
SQ SEQUENCE 184 AA; 20480 MW; 21454B241C717853 CRC64;

Query Match 93.2%; Score 109; DB 2; Length 184;
Best Local Similarity 95.5%; Pred. No. 4.6e-09;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KAEVQIARKLQCIADQFHLHV 22
Db 128 QAEVQIARKLQCIADQFHLHV 149

RESULT 5
AAH69505 PRELIMINARY; PRT; 184 AA.
AC AAH69505;
DT 20-MAY-2004 (TrEMBLrel. 27, Created)

```

DT 20-MAY-2004 (TrEMBLrel. 27, Last sequence update)
 DE Bcl2 modifying factor.
 GN BMF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PCR rescued clones;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler N.K.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshlyuk S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra W.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PCR rescued clones;
 RA Strausberg R.;
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC069505; AAH69505.1; -.
 SQ SEQUENCE 184 AA; 20477 MW; 201787489662846 CRC64;

 Query Match 93.2%; Score 109; DB 2; Length 184;
 Best Local Similarity 95.5%; Pred. No. 4.6e-09;
 Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

 QY 1 KAEVQIARKLQCIADQFHLHV 22
 :|||||
 DB 128 QAEVQIARKLQCIADQFHLHV 149
 :|||||

 RESULT 6
 Q9H7K7 PRELIMINARY; PRT; 204 AA.
 AC Q9H7K7;
 DT 01-WAR-2001 (TrEMBLrel. 16, Created)
 DT 01-WAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-WAR-2001 (TrEMBLrel. 16, Last sequence update)
 DE FLJ00065 protein (fragment).
 GN Name=FLJ00065;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RX PubMed=14702039;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Oobayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
 RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
 RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,

RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
 RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
 RA Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y., Ishida S.,
 RA Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hotuta T., Kusano J.,
 RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
 RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
 RA Mutsaers H., Yuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yashikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Koniyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani K., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Komatsu T.,
 RA Tazushi T., Cyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 DR EMBL; AK024472; BAB15762.1; -.
 FT NON TER 1
 SQ SEQUENCE 204 AA; 22708 MW; 5CC72C92664A688C CRC64;

 Query Match 93.2%; Score 109; DB 2; Length 204;
 Best Local Similarity 95.5%; Pred. No. 5.2e-09;
 Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

 QY 1 KAEVQIARKLQCIADQFHLHV 22
 :|||||
 DB 148 QAEVQIARKLQCIADQFHLHV 169
 :|||||

 RESULT 7
 Q6P9F6 PRELIMINARY; PRT; ~ 261 AA.
 ID Q6P9F6
 AC Q6P9F6;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE BMF protein (fragment) (fragment).
 GN Name=BMF;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshlyuk S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra W.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";

```

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC060783; AAH60783.1; -
DR EMBL; BC070043; AAH70043.1; -
FT NON TER 1
SQ SEQUENCE 261 AA; 29186 MW; FF8C6F214B0C24EB CRC64;

Query Match 93.2%; Score 109; DB 2; Length 261;
Best Local Similarity 95.5%; Pred. No. 6.7e-09;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAEVOIARKLQCIADQFHLHV 22
DB 205 QAEVOIARKLQCIADQFHLHV 226

RESULT 9
AAH70043 PRELIMINARY; PRT; 261 AA.
AC AAH70043;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Touchman J.W., Green E.D., Dickson M.C., Butterfield Y.S.,
RA Blakesley R.W., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC070043; AAH70043.1; -
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 261 AA; 29186 MW; FF8C6F214B0C24EB CRC64;

Query Match 93.2%; Score 109; DB 2; Length 261;
Best Local Similarity 95.5%; Pred. No. 6.7e-09;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAEVOIARKLQCIADQFHLHV 22
DB 205 QAEVOIARKLQCIADQFHLHV 226

RESULT 10
Q8K589 PRELIMINARY; PRT; 185 AA.
ID Q8K589
AC Q8K589;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE BMF protein.
GN Name=Bmf.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```



```

OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=22672518; PubMed=12787069;
RA Itoh T., Itoh A., Pleasure D.;
RT "Bcl-2-related protein family gene expression during oligodendroglial
RT differentiation.";
RL J. Neurochem. 85:1500-1512(2003).
DR EMBL: AF506761; AAM28890.1; -.
SQ SEQUENCE 185 AA; 20710 MW; A8F8D9FF10AD15B4 CRC64;

Query Match 91.5%; Score 107; DB 2; Length 185;
Best Local Similarity 90.9%; Pred. No. 9.7e-09;
Matches 20; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KAEVQIARKLQCIADQPHRLH 22
Db 129 RAEVQIARKLQCIADQPHRLH 150

RESULT 11
Q91Z99 PRELIMINARY; PRT; 185 AA.
AC Q91Z99
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Bcl-2 modifying factor (Mus musculus 0 day neonate thymus cDNA, RIKEN
DE full-length enriched library, clone:A430110F10 product:BCL-2 MODIFYING
DE FACTOR homolog.).
GN Name=Bmf;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=21432178; PubMed=11546872;
RA Futhalakath H., Villunger A., O'Reilly L.A., Beaumont J.G.,
RA Coutlas L., Cheney R.E., Huang D.C., Strasser A.;
RT "Bmf, a proapoptotic BH3-only protein regulated by interaction with
RT the myosin V actin motor complex, activated by anokis.";
RL Science 293:1829-1832(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RA The FANTOM Consortium;
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,

```

```

RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohata E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RA Agachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Haragaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito K., Saitoh K., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tegami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY029253; AAK38747.1; -.
DR EMBL: AK040822; BAC30649.1; -.
DR MGD; MGI:2176433; Bmf.
DR GO; GO:0015629; C:actin cytoskeleton; IDA.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0042981; P:regulation of apoptosis; IDA.
SQ SEQUENCE 185 AA; 20682 MW; DAE5EA080F3EA06B CRC64;

Query Match 90.6%; Score 106; DB 2; Length 185;
Best Local Similarity 95.2%; Pred. No. 1.4e-08;
Matches 20; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KAEVQIARKLQCIADQPHRLH 21
Db 129 RAEVQIARKLQCIADQPHRLH 149

RESULT 12
Q8BUKO PRELIMINARY; PRT; 271 AA.
AC Q8BUKO;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 13 days embryo heart cDNA, RIKEN full-length enriched
DE library, clone:D330026D15 product:BCL-2 MODIFYING FACTOR homolog.
GN Name=Bmf;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]

```

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RA The FANTOM Consortium
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RX MEDLINE=20493374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
Kanno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RA "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Katsunai T., Tashiro H., Itoh M.,
Kanno H., Akiyama J., Nishi K., Katsunai T., Nishine T., Harada A.,
Sumi N., Ishii Y., Nakamura S., Hazama M., Nakamura K.,
Yamamoto R., Matsumoto H., Sakauchi S., Ikegami T., Kashiwagi K.,
Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataniki M.,
Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RA "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,
Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RA Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AK084658; BAC39243.1; -;
DR MGD; MGI:2176433; Bmi;
DR GO; GO:0015629; C:actin cytoskeleton; IDA.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0042981; P:regulation of apoptosis; IDA.
SQ SEQUENCE 271 AA; 30071 MW; D4E57B0AC0998E98 CRC64;
Query Match 90.8%; Score 106; DB 2; Length 271;
Best Local Similarity 95.2%; Pred. No. 2.1e-08;
Matches 20; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 KAEVQIARKLQCIADQFHLHL 21
DB 215 RAEVQIARKLQCIADQFHLHL 235
RESULT 13
Q938D4 PRELIMINARY; PRT; 269 AA.
ID Q938D4
AC Q938D4;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative oxidoreductase.
OS Mycobacterium smegmatis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1772;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=mc2_155;
RA Derbyshire K.M., Parsons L.M., DeVost J.;
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
(GDR) family.
CC EMBL; AY054120; AAL17926.1; -;
DR HSSP; P97852; 1GZ6.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002198; ADH short.
DR InterPro; IPR002347; Adh short_C2.
DR Pfam; PF00106; adh short; 1.
DR PRINTS; PR00081; GDRDH.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
KW Oxidoreductase.
SQ SEQUENCE 269 AA; 28106 MW; E7F32C59CA4CB081 CRC64;
Query Match 43.6%; Score 51; DB 2; Length 269;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 6 IARKLQCIADQFHLHL 23
DB 91 VARSVQAIADFGFRDIL 108
RESULT 14
Q7U7Z4 PRELIMINARY; PRT; 325 AA.
AC Q7U7Z4
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Ordered locus names=SYNW0834;
OS Synechococcus sp. (strain WH8102).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=84589;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825697; PubMed=12917641; DOI=10.1038/nature01943;
RA Palenik B., Brahmsha B., Larimer F.W., Land M.L., Hauser L.,
Chain P., Lamerdin J.E., Regala W., Allen E.B., McCarran J.,
Paulsen I.T., Dufresne A., Partensky F., Webb E.A., Waterbury J.;
RT "The genome of a motile marine Synechococcus.";
RL Nature 424:1037-1042(2003).
DR EMBL; BX569691; CA807349.1; -;
DR InterPro; IPR010349; Asparaginase_II.
DR Pfam; PF06089; Asparaginase_II; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 325 AA; 34588 MW; 72FD0DA9C0C81895 CRC64;
Query Match 41.9%; Score 49; DB 2; Length 325;
Best Local Similarity 36.4%; Pred. No. 28;
Matches 8; Conservative 6; Mismatches 8; Indels 0; Gaps 0;
QY 2 AEVQIARKLQCIADQFHLHL 23
DB 16 APLEVTKRGCIASVHRVAV 37
RESULT 15

Q72Q79
 ID Q72Q79 PRELIMINARY; PRT; 351 AA.
 AC Q72Q79;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Fructose-bisphosphate aldolase.
 GN OrderedLocusNames=LIC12233;
 OS Leptospira interrogans (serogroup Icterohaemorrhagiae / serovar
 OC Copenhagen).
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
 OX NCBI_TaxID=44275;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Fiocruz Li-130;
 RX PubMed=15028702;
 RA Nascimento A.L.T.O., Ko A.I., Martins E.A.L., Monteiro-Vitorello C.B.,
 RA Ho P.L., Haake D.A., Verjovski-Almeida S., Hartskeerl R.A.,
 RA Marques M.V., Oliveira M.C., Menck C.F.M., Leite L.C.C., Carrer H.,
 RA Coutinho L.L., Degraive W.M., Dellagostin O.A., El-Dorri H.,
 RA Ferro E.S., Ferro M.I.T., Furlan L.R., Gamberini M., Giglioti E.A.,
 RA Goes-Neto A., Goldman G.H., Goldman M.H.S., Harakava R.,
 RA Jeronimo S.M.B., Junqueira-de-Azevedo I.L.M., Kimura E.T.,
 RA Kuramae E.E., Lemos E.G.M., Lemos M.V.F., Marino C.L., Nunes L.R.,
 RA de Oliveira R.C., Pereira G.G., Reis M.S., Schriefer A.,
 RA Siqueira W.J., Sommer P., Tsai S.M., Simpson A.J.G., Ferro J.A.,
 RA Camargo L.E.A., Kitajima J.P., Setubal J.C., Van Sluys M.A.;
 RT "Comparative genomics of two Leptospira interrogans serovars reveals
 RT novel insights into physiology and pathogenesis.";
 RL J. Bacteriol. 186:2164-2172(2004).
 DR EMBL; AF017295; AAS70805.1; -.
 KW Complete proteome.
 SQ SEQUENCE 351 AA; 37977 MW; EE7A6D95E432F812 CRC64;
 Query Match 41.0%; Score 48; DB 2; Length 351;
 Best Local Similarity 52.9%; Pred. No. 44;
 Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 QY 7 ARKLCIADQFRLHVL 23
 Db :|:|:|:|:|:|
 177 SRQIQEISDAFHRAHEL 193

Search completed: November 10, 2004, 15:53:41
 Job time : 40.6285 secs

This page blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:55:47 ; Search time 10.1779 Seconds
(without alignments)
143.349 Million cell updates/sec

Title: US-10-092-750-227

Perfect score: 110

Sequence: 1 MGDVVGFIDELEGAVSDLHRL 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/1/aa/5A.COMB.pgp: *
2: /cgn2_6/ptodata/1/aa/5B.COMB.pgp: *
3: /cgn2_6/ptodata/1/aa/6A.COMB.pgp: *
4: /cgn2_6/ptodata/1/aa/6B.COMB.pgp: *
5: /cgn2_6/ptodata/1/aa/PCTUS.COMB.pgp: *
6: /cgn2_6/ptodata/1/aa/backfiles1.pgp: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query | | DB | ID | Description |
|------------|-------|--------------|------|----|----------------------|
| | Score | Match Length | | | |
| 1 | 50.5 | 45.9 | 1824 | 2 | US-08-680-327-3 |
| 2 | 50.5 | 45.9 | 1824 | 3 | US-09-228-246-2 |
| 3 | 47 | 42.7 | 343 | 4 | US-09-377-285B-28 |
| 4 | 47 | 42.7 | 354 | 4 | US-09-377-285B-30 |
| 5 | 45 | 40.9 | 159 | 4 | US-09-489-039A-10657 |
| 6 | 45 | 40.9 | 319 | 4 | US-09-252-991A-18150 |
| 7 | 44.5 | 40.5 | 178 | 4 | US-09-252-991A-21443 |
| 8 | 44.5 | 40.5 | 317 | 3 | US-09-199-637A-433 |
| 9 | 44.5 | 40.5 | 377 | 4 | US-09-252-991A-19827 |
| 10 | 44 | 40.0 | 154 | 4 | US-09-621-976-4100 |
| 11 | 44 | 40.0 | 249 | 4 | US-09-673-395A-586 |
| 12 | 44 | 40.0 | 309 | 4 | US-09-248-796A-14974 |
| 13 | 44 | 40.0 | 346 | 3 | US-09-199-637A-313 |
| 14 | 44 | 40.0 | 346 | 4 | US-09-252-991A-21432 |
| 15 | 44 | 40.0 | 538 | 4 | US-09-134-000C-5458 |
| 16 | 44 | 40.0 | 5087 | 3 | US-09-144-085-1 |
| 17 | 44 | 40.0 | 6095 | 3 | US-09-144-085-2 |
| 18 | 43.5 | 39.5 | 141 | 2 | US-08-627-173-21 |
| 19 | 43.5 | 39.5 | 141 | 2 | US-08-535-882A-21 |
| 20 | 43.5 | 39.5 | 141 | 3 | US-08-316-424A-7 |
| 21 | 43.5 | 39.5 | 141 | 3 | US-09-005-546-21 |
| 22 | 43.5 | 39.5 | 141 | 4 | US-08-477-669-7 |
| 23 | 43.5 | 39.5 | 141 | 4 | US-10-128-581-26 |
| 24 | 43.5 | 39.5 | 141 | 4 | US-08-832-443-21 |
| 25 | 43.5 | 39.5 | 1262 | 3 | US-09-357-251-33 |
| 26 | 43.5 | 39.5 | 1266 | 1 | US-08-468-557-4 |
| 27 | 43.5 | 39.5 | 1266 | 3 | US-09-357-251-32 |

| | | | | | | |
|----|------|------|-----|---|----------------------|-------------------|
| 28 | 43 | 39.1 | 100 | 4 | US-09-248-796A-23040 | Sequence 23040, A |
| 29 | 43 | 39.1 | 128 | 4 | US-09-248-796A-27548 | Sequence 27548, A |
| 30 | 43 | 39.1 | 141 | 1 | US-08-240-712-18 | Sequence 18, Appl |
| 31 | 43 | 39.1 | 141 | 1 | US-08-443-890-18 | Sequence 18, Appl |
| 32 | 43 | 39.1 | 141 | 2 | US-08-484-686B-65 | Sequence 65, Appl |
| 33 | 43 | 39.1 | 141 | 3 | US-08-463-160B-66 | Sequence 66, Appl |
| 34 | 43 | 39.1 | 141 | 3 | US-09-058-563-18 | Sequence 18, Appl |
| 35 | 43 | 39.1 | 141 | 5 | PCT-US92-09752-18 | Sequence 18, Appl |
| 36 | 43 | 39.1 | 349 | 3 | US-09-161-241-8 | Sequence 8, Appl |
| 37 | 43 | 39.1 | 473 | 4 | US-09-328-352-4658 | Sequence 4658, Ap |
| 38 | 43 | 39.1 | 522 | 4 | US-09-583-110-5045 | Sequence 5045, Ap |
| 39 | 43 | 39.1 | 546 | 4 | US-09-252-991A-22048 | Sequence 22048, A |
| 40 | 42.5 | 38.6 | 146 | 1 | US-08-240-712-21 | Sequence 21, Appl |
| 41 | 42.5 | 38.6 | 146 | 1 | US-08-240-712-22 | Sequence 22, Appl |
| 42 | 42.5 | 38.6 | 146 | 1 | US-08-443-890-21 | Sequence 21, Appl |
| 43 | 42.5 | 38.6 | 146 | 1 | US-08-443-890-22 | Sequence 22, Appl |
| 44 | 42.5 | 38.6 | 146 | 2 | US-08-484-686B-67 | Sequence 67, Appl |
| 45 | 42.5 | 38.6 | 146 | 3 | US-08-463-160B-68 | Sequence 68, Appl |

ALIGNMENTS

RESULT 1
US-08-680-327-3
; Sequence 3, Application US/08680327
; Patent No. 5859321
; GENERAL INFORMATION:
; APPLICANT: Staskawicz, Brian S., Oldroyd, Giles Edward,
; APPLICANT: Salmeron, John M., Rommens, Cains
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PLANT
; TITLE OF INVENTION: PATHOGEN RESISTANCE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klarquist Sparkman Campbell Leigh &
; ADDRESSEE: Whinston
; STREET: One World Trade Center
; STREET: 121 S.W. Salmon Street
; STREET: Suite 1600
; CITY: Portland
; STATE: Oregon
; COUNTRY: United States of America
; ZIP: 97204
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Disk, 3-1/2 inch
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/680,327
; FILING DATE: July 11, 1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/310,912
; FILING DATE: September 22, 1994
; CLASSIFICATION: 800
; APPLICATION NUMBER: 08/227,360
; FILING DATE: April 13, 1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Alan. E.
; REGISTRATION NUMBER: 35,123
; REFERENCE/DOCKET NUMBER: 5151-45038
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (503) 226-7391
; TELEFAX: (503) 228-9446
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1824 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-680-327-3


```

Best Local Similarity 64.3%; Pred No. 12;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 GFIDELEGAVSLH 19
   ||: ||: |||
Db 126 GFLRELDPEVSLH 139

RESULT 6
US-09-252-991A-18150
; Sequence 18150, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18150
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18150

Query Match 40.9%; Score 45; DB 4; Length 319;
Best Local Similarity 57.1%; Pred. No. 26;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 7 FIDELEGAVSLHR 20
   ||: ||: |||
Db 296 FVDRLDGAVGEYR 309

RESULT 7
US-09-252-991A-21443
; Sequence 21443, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21443
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21443

Query Match 40.5%; Score 44.5; DB 4; Length 178;
Best Local Similarity 32.3%; Pred. No. 16;
Matches 10; Conservative 5; Mismatches 5; Indels 11; Gaps 1;

QY 3 DVVGFIDELEGAVS-----DLHRL 22
   ||: ||: |||
Db 116 DIEGTVDLDGAIHAGAAAGIGFIDLHGL 146

RESULT 8
US-09-199-637A-433
; Sequence 433, Application US/09199637A

```

```

; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Tscengalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 433
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-433

Query Match 40.5%; Score 44.5; DB 3; Length 317;
Best Local Similarity 32.3%; Pred. No. 31;
Matches 10; Conservative 5; Mismatches 5; Indels 11; Gaps 1;

QY 3 DVVGFIDELEGAVS-----DLHRL 22
   ||: ||: |||
Db 255 DIEGTVDLDGAIHAGAAAGIGFIDLHGL 285

RESULT 9
US-09-252-991A-19827
; Sequence 19827, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19827
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19827

Query Match 40.5%; Score 44.5; DB 4; Length 377;
Best Local Similarity 50.0%; Pred. No. 39;
Matches 10; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

QY 1 MGDVVGFIIDELEGA-VSDLH 19
   ||: ||: |||
Db 239 LGQVVGQFVDGVPGLVADAH 258

RESULT 10
US-09-621-976-4100
; Sequence 4100, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.

```

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2

; CURRENT APPLICATION NUMBER: US/09/621,976

; CURRENT FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 19335

; SOFTWARE: Patent.pm

; SEQ ID NO 4100

; LENGTH: 154

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SIGNAL

; LOCATION: -147...-1

US-09-621-976-4100

Query Match 40.0%; Score 44; DB 4; Length 154;

Best Local Similarity 45.0%; Pred. No. 16;

Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 3 DVVGFIDELEGAVSDLHRL 22

Db 108 DIASDDLEAARNELHSAL 127

RESULT 11

US-09-673-395A-586

; Sequence 586, Application US/09673395A

; Patent No. 6620923

; GENERAL INFORMATION:

; APPLICANT: SPECHT, THOMAS

; APPLICANT: HINZMANN, BERND

; APPLICANT: SCHMITT, ARMIN

; APPLICANT: PILARSKY, CHRISTIAN

; APPLICANT: DAHL, EDGAR

; APPLICANT: ROSENTHAL, ANDRE

; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE

; FILE REFERENCE: ALBRE-12

; CURRENT APPLICATION NUMBER: US/09/673,395A

; CURRENT FILING DATE: 2000-10-17

; NUMBER OF SEQ ID NOS: 637

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 586

; LENGTH: 249

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-673-395A-586

Query Match 40.0%; Score 44; DB 4; Length 249;

Best Local Similarity 68.4%; Pred. No. 29;

Matches 13; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

QY 4 VVGFIDELEG--AVSDLHR 20

Db 62 VVGVGDELEGAVSGDHR 80

RESULT 12

US-09-248-796A-14974

; Sequence 14974, Application US/09248796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

; FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS

; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 14974

; LENGTH: 309

; TYPE: PRT

; ORGANISM: Candida albicans

US-09-248-796A-14974

Query Match 40.0%; Score 44; DB 4; Length 309;

Best Local Similarity 25.8%; Pred. No. 37;

Matches 10; Conservative 5; Mismatches 4; Indels 20; Gaps 1;

QY 1 MGDVVG-----FIDELEGAVSDLH 19

Db 103 MGDIVGTLLHETLNLYSAHDSKFYKFLDKJRSYDDIH 141

RESULT 13

US-09-199-637A-313

; Sequence 313, Application US/09199637A

; Patent No. 6355411

; GENERAL INFORMATION:

; APPLICANT: Ausubel, Frederick

; APPLICANT: Goodman, Howard M.

; APPLICANT: Rahme, Laurence G.

; APPLICANT: Mahajan-Miklos, Shalina

; APPLICANT: Tan, Man-Wah

; APPLICANT: Cao, Hui

; APPLICANT: Drenkard, Eliana

; APPLICANT: Tsongalis, John

; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID

; FILE REFERENCE: SEQUENCES AND USES THEREOF

; CURRENT APPLICATION NUMBER: US/09/199,637A

; CURRENT FILING DATE: 1998-11-25

; PRIOR APPLICATION NUMBER: 60/066,517

; PRIOR FILING DATE: 1997-11-25

; NUMBER OF SEQ ID NOS: 437

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 313

; LENGTH: 346

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-199-637A-313

Query Match 40.0%; Score 44; DB 3; Length 346;

Best Local Similarity 42.9%; Pred. No. 42;

Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 MGDVVGFIDELEGAVSDLHRA 21

Db 273 LGFLLGLQDLRAAVDDVARA 293

RESULT 14

US-09-252-991A-21432

; Sequence 21432, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 21432

; LENGTH: 346

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-21432

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 02:43:24 ; Search time 53.1601 Seconds
(without alignments)
146.426 Million cell updates/sec

Title: US-10-092-750-227
Perfect score: 110
Sequence: 1 MGDVVGFIDELEGAUSDHRLAL 22

Scoring table: BLOSUM62
Gapext 0.5

Searched: 1568699 seqs, 353819137 residues

Total number of hits satisfying chosen parameters: 1568699

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09D_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 110 | 100.0 | 22 | 14 | US-10-092-750-227 |
| 2 | 61 | 55.5 | 22 | 14 | US-10-092-750-61 |
| 3 | 50 | 45.5 | 225 | 9 | US-09-738-626-5107 |
| 4 | 50 | 45.5 | 232 | 15 | US-10-425-114-37071 |
| 5 | 50 | 45.5 | 255 | 17 | US-10-425-115-219891 |
| 6 | 49 | 44.5 | 331 | 14 | US-10-369-493-14158 |
| 7 | 49 | 44.5 | 478 | 14 | US-10-369-493-11767 |
| 8 | 49 | 44.5 | 490 | 14 | US-10-369-493-14503 |
| 9 | 49 | 44.5 | 490 | 14 | US-10-369-493-14898 |
| 10 | 48 | 43.6 | 233 | 15 | US-10-424-599-149605 |
| 11 | 48 | 43.6 | 316 | 17 | US-10-425-115-291868 |
| 12 | 48 | 43.6 | 450 | 17 | US-10-739-930-8810 |
| 13 | 48 | 43.6 | 494 | 10 | US-09-934-455-26 |

Query Match 100.0%; Score 110; DB 14; Length 22;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGDVVGFIDELEGAUSDHRLAL 22
Db 1 MGDVVGFIDELEGAUSDHRLAL 22

RESULT 2
US-10-092-750-61
; Sequence 61, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:

ALIGNMENTS

RESULT 1

US-10-092-750-227
; Sequence 227, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION: Philip W.
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 227
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-227

| | | | | | | |
|----|------|------|------|----|----------------------|--------------------|
| 14 | 48 | 43.6 | 494 | 14 | US-10-286-264-56 | Sequence 56, Appl |
| 15 | 48 | 43.6 | 494 | 14 | US-10-225-068-148 | Sequence 148, App |
| 16 | 48 | 43.6 | 494 | 14 | US-10-225-066A-228 | Sequence 228, App |
| 17 | 48 | 43.6 | 494 | 14 | US-10-302-267-90 | Sequence 90, Appl |
| 18 | 48 | 43.6 | 494 | 14 | US-10-425-913-2 | Sequence 2, Appl |
| 19 | 48 | 43.6 | 494 | 15 | US-10-374-780A-216 | Sequence 216, App |
| 20 | 48 | 43.6 | 494 | 15 | US-10-412-699B-714 | Sequence 714, App |
| 21 | 48 | 43.6 | 3173 | 10 | US-09-918-715-218 | Sequence 218, App |
| 22 | 48 | 43.6 | 3173 | 17 | US-10-474-794-218 | Sequence 218, App |
| 23 | 48 | 43.6 | 3176 | 15 | US-10-372-683-20 | Sequence 20, Appl |
| 24 | 48 | 43.6 | 3176 | 16 | US-10-734-564-91 | Sequence 91, Appl |
| 25 | 47.5 | 43.2 | 330 | 16 | US-10-437-963-113212 | Sequence 113212 |
| 26 | 47 | 42.7 | 253 | 15 | US-10-425-114-40181 | Sequence 40181, A |
| 27 | 47 | 42.7 | 283 | 15 | US-10-424-599-283577 | Sequence 283577 |
| 28 | 47 | 42.7 | 290 | 16 | US-10-437-963-132914 | Sequence 132914 |
| 29 | 47 | 42.7 | 343 | 14 | US-10-192-381-28 | Sequence 28, Appl |
| 30 | 47 | 42.7 | 354 | 14 | US-10-192-381-30 | Sequence 30, Appl |
| 31 | 47 | 42.7 | 412 | 15 | US-10-374-780A-1378 | Sequence 1378, Ap |
| 32 | 47 | 42.7 | 412 | 15 | US-10-412-699B-1499 | Sequence 1499, Ap |
| 33 | 47 | 42.7 | 412 | 16 | US-10-437-963-103217 | Sequence 103217 |
| 34 | 47 | 42.7 | 450 | 14 | US-10-302-267-96 | Sequence 96, Appl |
| 35 | 47 | 42.7 | 450 | 15 | US-10-374-780A-2128 | Sequence 2128, Ap |
| 36 | 47 | 42.7 | 450 | 15 | US-10-412-699B-934 | Sequence 934, App |
| 37 | 47 | 42.7 | 450 | 15 | US-10-412-699B-1870 | Sequence 1870, App |
| 38 | 47 | 42.7 | 828 | 14 | US-10-425-913-3 | Sequence 3, Appl |
| 39 | 46 | 41.8 | 143 | 17 | US-10-425-115-311287 | Sequence 311287 |
| 40 | 46 | 41.8 | 153 | 15 | US-10-424-599-275789 | Sequence 275789 |
| 41 | 46 | 41.8 | 173 | 15 | US-10-424-599-170417 | Sequence 170417 |
| 42 | 46 | 41.8 | 338 | 16 | US-10-437-963-163935 | Sequence 163935 |
| 43 | 46 | 41.8 | 373 | 16 | US-10-437-963-157518 | Sequence 157518 |
| 44 | 46 | 41.8 | 1077 | 14 | US-10-128-714-3179 | Sequence 3179, Ap |
| 45 | 46 | 41.8 | 1077 | 14 | US-10-128-714-8179 | Sequence 8179, Ap |

; APPLICANT: Hammond, Philip W.
 ; APPLICANT: Alpin, Julia
 ; APPLICANT: Wright, Martin C.
 ; TITLE OF INVENTION: Polypeptides Interactive with BCL-XL
 ; FILE REFERENCE: 50036/050002
 ; CURRENT APPLICATION NUMBER: US/10/092,750
 ; CURRENT FILING DATE: 2002-03-07
 ; PRIOR APPLICATION NUMBER: US 60/274,526
 ; PRIOR FILING DATE: 2001-03-08
 ; NUMBER OF SEQ ID NOS: 253
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 61
 ; LENGTH: 22
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; ORGANISM: Zea mays
 ; OTHER INFORMATION: Clone ID: LIB3180-055-A2_FLI.pep
 ; US-10-092-750-61

Query Match 55.5%; Score 61; DB 14; Length 22;
 Best Local Similarity 75.0%; Pred. No. 0.019;
 Matches 15; Conservative 0; Mismatches 0; Gaps 0; Indels 5;

QY 3 DVVGFIDELEGAVSDLHRL 22
 DB 3 DVVGFIDELEGAVSDLHRL 22

RESULT 3
 US-09-738-626-5107
 ; Sequence 5107, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; CURRENT FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 5107
 ; LENGTH: 225
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 ; ORGANISM: Zea mays
 ; OTHER INFORMATION: Clone ID: MRT4577_132128C.1.pep
 ; US-09-738-626-5107

Query Match 45.5%; Score 50; DB 9; Length 225;
 Best Local Similarity 47.8%; Pred. No. 15;
 Matches 11; Conservative 7; Mismatches 3; Gaps 1; Indels 3;

QY 2 GDVVGFI--DELEGAVSDLHRL 22
 DB 145 GNVNGFLAEDGRPALANTHRL 167

RESULT 4
 US-10-425-114-37071
 ; Sequence 37071, Application US/10425114
 ; Publication No. US20040034886A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; PLANTS WITH IMPROVED PROPERTIES

; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaska, Jack E.
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 37071
 ; LENGTH: 232
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: LIB3180-055-A2_FLI.pep
 ; US-10-425-114-37071

Query Match 45.5%; Score 50; DB 15; Length 232;
 Best Local Similarity 40.9%; Pred. No. 15;
 Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 MGDVVGFIDELEGAVSDLHRL 22
 DB 80 LGDAIEYKELLQRLSDHNL 101

RESULT 5
 US-10-425-115-219891
 ; Sequence 219891, Application US/10425115
 ; Publication No. US20040214272A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants
 ; FILE REFERENCE: 38-21(53222)B
 ; CURRENT APPLICATION NUMBER: US/10/425,115
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 369326
 ; SEQ ID NO 219891
 ; LENGTH: 255
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: MRT4577_132128C.1.pep
 ; US-10-425-115-219891

Query Match 45.5%; Score 50; DB 17; Length 255;
 Best Local Similarity 40.9%; Pred. No. 17;
 Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 MGDVVGFIDELEGAVSDLHRL 22
 DB 103 LGDAIEYKELLQRLSDHNL 124

RESULT 6
 US-10-369-493-14158
 ; Sequence 14158, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; PLANTS WITH IMPROVED PROPERTIES

```
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14158
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-14158

Query Match          44.5%; Score 49; DB 14; Length 331;
Best Local Similarity 56.2%; Pred. No. 33;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy      2 GDVVGFIDELEGAUSD 17
Db      288 GPKVGYLEELEGGVSN 303

RESULT 7
US-10-369-493-11767
; Sequence 11767, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 11767
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-11767

Query Match          44.5%; Score 49; DB 14; Length 478;
Best Local Similarity 56.2%; Pred. No. 50;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy      2 GDVVGFIDELEGAUSD 17
Db      288 GPKVGYLEELEGGVSN 303

RESULT 8
US-10-369-493-14503
; Sequence 14503, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14503

; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14158
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-14158

Query Match          44.5%; Score 49; DB 14; Length 331;
Best Local Similarity 56.2%; Pred. No. 33;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy      2 GDVVGFIDELEGAUSD 17
Db      288 GPKVGYLEELEGGVSN 303

RESULT 9
US-10-369-493-14898
; Sequence 14898, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14898
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-14898

Query Match          44.5%; Score 49; DB 14; Length 490;
Best Local Similarity 56.2%; Pred. No. 51;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy      2 GDVVGFIDELEGAUSD 17
Db      299 GPKVGYLEELEGGVSN 314

RESULT 10
US-10-424-599-149605
; Sequence 149605, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 149605
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_106115C.1.pap
US-10-424-599-149605

Query Match          43.6%; Score 48; DB 15; Length 233;
Best Local Similarity 36.4%; Pred. No. 31;
Matches 8; Conservative 6; Mismatches 8; Indels 0; Gaps 0;
```


Best Local Similarity 36.4%; Pred. No. 74;
Matches 8; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 MGDVVGFIDELEGAVSDHLRAL 22
DB 342 LGDAIDYKELQLQRINDLHNL 363

RESULT 15
US-10-225-068-148
; Sequence 148, Application US/10225068
; Publication No. US20030217383A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Adam, Luc J.
; APPLICANT: Dubell, Arnold T.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Broun, Pierre E.
; TITLE OF INVENTION: STRESS-RELATED POLYNUCLEOTIDES AND
; TITLE OF INVENTION: POLYPEPTIDES IN PLANTS
; FILE REFERENCE: 514442002040
; CURRENT APPLICATION NUMBER: US/10/225,068
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 246
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 148
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (308)...(359)
; OTHER INFORMATION: Conserved domain
US-10-225-068-148

Query Match 43.5%; Score 48; DB 14; Length 494;
Best Local Similarity 36.4%; Pred. No. 74;
Matches 8; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 MGDVVGFIDELEGAVSDHLRAL 22
DB 342 LGDAIDYKELQLQRINDLHNL 363

Search completed: November 11, 2004, 07:41:47
Job time : 53.1601 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:52:32 ; Search time 7.51601 Seconds
(without alignments)
281.634 Million cell updates/sec

Title: US-10-092-750-227
Perfect score: 110
Sequence: 1 MGDDVGFIDELEGAVSDLHRLAL 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 segs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|------------------------|
| 1 | 51 | 46.4 | 197 | 2 A87286 | conserved hypothetical |
| 2 | 50.5 | 45.9 | 1824 | 2 T07589 | disease resistance |
| 3 | 50 | 45.5 | 523 | 2 B75467 | phosphoribosylamin |
| 4 | 49 | 44.5 | 391 | 2 T04609 | hypothetical prote |
| 5 | 49 | 44.5 | 491 | 2 AC2650 | glucose-6-phosphat |
| 6 | 49 | 44.5 | 503 | 2 B97432 | glucose-6-phosphat |
| 7 | 48.5 | 44.1 | 141 | 1 HATJQA | hemoglobin alpha-A |
| 8 | 48.5 | 44.1 | 144 | 1 HATRL1 | hemoglobin I alpha |
| 9 | 48 | 43.6 | 141 | 1 HGWMP | hemoglobin alpha c |
| 10 | 48 | 43.6 | 3176 | 2 CGHU3A | collagen alpha 3(V |
| 11 | 47 | 42.7 | 218 | 2 A47486 | glutathione transf |
| 12 | 47 | 42.7 | 218 | 2 S32425 | glutathione transf |
| 13 | 47 | 42.7 | 388 | 2 T01024 | hypothetical prote |
| 14 | 47 | 42.7 | 450 | 2 C82622 | Fl3k23.12 protein |
| 15 | 47 | 42.7 | 528 | 2 AG2288 | GRP-binding protei |
| 16 | 47 | 42.7 | 1200 | 2 C96025 | hypothetical expor |
| 17 | 46.5 | 42.3 | 141 | 1 HAFXB | hemoglobin alpha c |
| 18 | 46.5 | 42.3 | 141 | 1 S59496 | hemoglobin alpha c |
| 19 | 46.5 | 42.3 | 143 | 1 S04589 | hemoglobin alpha c |
| 20 | 46 | 41.8 | 187 | 2 F70164 | grpe protein - Lym |
| 21 | 46 | 41.8 | 233 | 2 C98188 | cytochrome c bindi |
| 22 | 46 | 41.8 | 233 | 2 AF3098 | hypothetical prote |
| 23 | 46 | 41.8 | 590 | 1 GBBPT4 | baseplate protein |
| 24 | 46 | 41.8 | 1106 | 1 A97819 | isoleucine-tRNA li |
| 25 | 45.5 | 41.4 | 141 | 1 HACO | hemoglobin alpha c |
| 26 | 45.5 | 41.4 | 141 | 2 A25357 | hemoglobin alpha c |
| 27 | 45.5 | 41.4 | 141 | 2 A23702 | hemoglobin alpha c |
| 28 | 45.5 | 41.4 | 141 | 2 S28934 | hemoglobin alpha c |
| 29 | 45.5 | 41.4 | 142 | 1 HAGY | hemoglobin alpha c |

| | | | | | |
|----|------|------|------|----------|--------------------|
| 30 | 45.5 | 41.4 | 143 | 1 HACA | hemoglobin alpha c |
| 31 | 45.5 | 41.4 | 1039 | 2 C64418 | isoleucine-tRNA li |
| 32 | 45 | 40.9 | 227 | 2 AH2123 | hypothetical prote |
| 33 | 45 | 40.9 | 248 | 2 F72262 | ubiquinone/menaqui |
| 34 | 45 | 40.9 | 300 | 2 AG2025 | cell-division prot |
| 35 | 45 | 40.9 | 477 | 2 JC4386 | adenyl cyclase-a |
| 36 | 45 | 40.9 | 511 | 2 A83729 | phosphoribosylamin |
| 37 | 45 | 40.9 | 632 | 2 E69407 | NADH oxidase (noxB |
| 38 | 45 | 40.9 | 632 | 2 G89306 | NADH oxidase (noxB |
| 39 | 45 | 40.9 | 642 | 2 T10861 | phaseolin G-box bi |
| 40 | 45 | 40.9 | 716 | 2 T03695 | delta 1 pyroline- |
| 41 | 45 | 40.9 | 744 | 2 D39135 | hypothetical prote |
| 42 | 44.5 | 40.5 | 140 | 2 D60515 | hemoglobin alpha-1 |
| 43 | 44.5 | 40.5 | 141 | 1 HABTF | hemoglobin alpha c |
| 44 | 44.5 | 40.5 | 141 | 1 HABTV | hemoglobin alpha c |
| 45 | 44.5 | 40.5 | 141 | 1 HAHOD | hemoglobin alpha c |

ALIGNMENTS

RESULT 1

A87286 conserved hypothetical protein CC0298 [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C/Accession: A87286
R/Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.E.
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:11259647
A/Accession: A87286
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-197 <SFO>
A/Cross-references: UNIPROT:Q9ABD1; GB:AE005673; NID:gl3421439; PIDN:AAK22285.1; GSPDB:GN
C/Genetics:
A/Gene: CC0298

Query Match 46.4%; Score 51; DB 2; Length 197;
Best Local Similarity 61.1%; Pred. No. 4;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Cy 5 VGFIDELEGAVSDLHRLAL 22
Db 145 IGFHLETEAVSDLHRRM 162

RESULT 2

T07589 disease resistance protein Prf - tomato
C/Species: Lycopersicon esculentum (tomato)
C/Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 09-Jul-2004
C/Accession: T07589
R/Salmeron, J.M.; Oldroyd, G.E.D.; Rommens, C.M.T.; Scofield, S.R.; Kim, H.; Lavelle, D.J.
Cell 86, 123-133, 1996
A/Title: Tomato Prf is a member of the leucine rich repeat class of plant disease resist
A/Reference number: Z16032; MUID:96291405; PMID:8689679
A/Accession: T07589
A/Status: preliminary; translated from GB/EMBL/DBJ

A/Residues: 1-1824 <SAL>
A/Cross-references: UNIPROT:Q96485; EMBL:U65391; NID:gl513143; PIDN:AAK49408.1; PID:gl51
A/Experimental source: strain Rio Grande
C/Genetics:
A/Gene: Prf

A/Map position: 5
A/Introns: 1437/2
C/Keywords: leucine zipper; nucleotide binding
Query Match 45.9%; Score 50.5; DB 2; Length 1824;


```
C:Keywords: blood; chromoprotein; erythrocyte; heme; heterotetramer; iron; metalloprotein
F:2-141/Domain: globin homology <GUB>
F:58/Binding site: oxygen (His) (distal axial ligand) #status predicted
F:87/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match      44.1%; Score 48.5; DB 1; Length 141;
Best Local Similarity 45.5%; Pred. No. 6.6; 4; Indels 3; Gaps 1;
Matches 10; Conservative 5; Mismatches 4; Indels 4; Gaps 1;

QY 1 MGDVVGFDLEBGA---VSDLH 19
Db 66 MGEAVNHLDDMAGALLKLSDLH 87

RESULT 8
HATRI
hemoglobin I alpha chain - rainbow trout (tentative sequence)
C:Species: Oncorhynchus mykiss (rainbow trout)
C:Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 12-Jul-2004
A:Accession: A02348
R:Bossa, F.; Barra, D.; Petruzzelli, R.; Martini, F.; Brunori, M.
Biochim. Biophys. Acta 536, 298-305, 1978
A:Title: Primary structure of hemoglobin from trout (Salmo irideus). Amino acid sequence
A:Reference number: A02348; MUID:79042280; PMID:708770
A:Accession: A02348
A:Molecule type: protein
A:Residues: 1-144 <BOS>
A:Cross-references: UNIPROT:P02019
A:Note: this hemoglobin has two more residues, 32-Asp and 33-Lys, than other fish hemog
C:Superfamily: globin; globin homology
C:Keywords: acetylated amino end; blood; chromoprotein; erythrocyte; heme; iron; metallo
F:2-144/Domain: globin homology <GLE>
F:1/Modified site: acetylated amino end (Ser) #status experimental
F:61/Binding site: oxygen (His) (distal axial ligand) #status predicted
F:90/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match      44.1%; Score 48.5; DB 1; Length 144;
Best Local Similarity 50.0%; Pred. No. 6.7;
Matches 11; Conservative 4; Mismatches 4; Indels 3; Gaps 1;

QY 1 MGDVVGFDLEBGA---AVSDLH 19
Db 69 IGRKAVGLMDLVGMSAUSDH 90

RESULT 9
HACMP
hemoglobin alpha chain - pale-throated sloth
C:Species: Bradypus tridactylus (pale-throated sloth)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 12-Jul-2004
C:Accession: S03998
R:Kleinschmidt, T.; Maerz, J.; Braunitzer, G.
Biol. Chem. Hoppe-Seyler 370, 303-308, 1989
A:Title: The primary structure of pale-throated three-toed sloth (Bradypus tridactylus,
A:Reference number: S03998; MUID:89335258; PMID:2757790
A:Accession: S03998
A:Molecule type: protein
A:Residues: 1-141 <KLE>
A:Cross-references: UNIPROT:PI4525
C:Superfamily: globin; globin homology
C:Keywords: blood; chromoprotein; erythrocyte; heme; heterotetramer; iron; metalloprotein
F:2-141/Domain: globin homology <GLB>
F:58/Binding site: oxygen (His) (distal axial ligand) #status predicted
F:87/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match      43.6%; Score 48; DB 1; Length 141;
Best Local Similarity 50.0%; Pred. No. 7.8;
Matches 11; Conservative 4; Mismatches 3; Indels 4; Gaps 1;

QY 1 MGD-----VVGFDLEBGAUSD 18
Db 62 VGDALTAVGHLDLPFGALS 83
```

```
RESULT 10
CGHU3A
collagen alpha 3(VI) chain precursor [validated] - human
N:Contents: collagen alpha 3(VI) chain, splice form A9/N10(-)
C:Species: Homo sapiens (man)
C:Date: 21-Nov-1993 #sequence_revision 12-Nov-1999 #text_change 16-Aug-2004
C:Accession: A59140; S13679; S24465; A57083; S28776; S00245; C31952; C29848; S26510; S48;
R:Chu, M.L.
submitted to GenBank, May 1998
A:Reference number: A59140
A:Accession: A59140
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-3176 <CHU>
A:Cross-references: UNIPROT:PI2111; GB:X52022; NID:G3127925; PIDN:CAA36267.1; PID:G312792;
R:Chu, M.L.; Zhang, R.Z.; Pan, T.; Stokes, D.; Conway, D.; Kuo, H.J.; Glanville, R.; Maye
EMBO J. 9, 385-393, 1990
A:Title: Mosaic structure of globular domains in the human type VI collagen alpha-3 chain
itors.
A:Reference number: S13679; MUID:90151612; PMID:1689238
A:Accession: S13679
A:Molecule type: mRNA
A:Residues: 1-30,237-313, 'CWV', 318-322, 'AR', 326-1815, 'FD', 1818-1819, 'ID', 1822-3176 <CHS>
A:Cross-references: EMBL:X52022; NID:G3127925
A:Accession: S24465
A:Molecule type: protein
A:Residues: 574-585,965-973, 'X', 975-976,1306-1325,1361-1377,1381-1401,1473-1506, 'X', 1508
-1982, 'X', 1964-1965,2018-2037,2374-2410,2445-2459,2466-2469, 'X', 2471-2474,2504-2508, 'X',
R:Zanussi, S.; Doliana, R.; Segat, D.; Bonaldo, P.; Colombatti, A.
J. Biol. Chem. 267, 24082-24089, 1992
A:Title: The human type VI collagen gene, mRNA and protein variants of the alpha3 chain
A:Reference number: S28776; MUID:93054780; PMID:1339440
A:Accession: A57083
A:Molecule type: DNA
A:Residues: 310-328 <ZAN>
A:Accession: S28776
A:Molecule type: mRNA
A:Residues: 32-126, 'AK', 128-136, 'L', 138-236 <ZAZ>
A:Cross-references: GB:S49432; NID:G260296; PIDN:AAB24261.1; PID:G260297
R:Chu, M.L.; Mann, K.; Deutzmann, R.; Pribula-Conway, D.; Hsu-Chen, C.C.; Bernard, M.P.;
Eur. J. Biochem. 168, 309-317, 1987
A:Title: Characterization of three constituent chains of collagen type VI by peptide seq
A:Reference number: S00126; MUID:88029444; PMID:3665927
A:Accession: S00245
A:Molecule type: mRNA; protein
A:Residues: 2024-2046; 2092-2156, 'R', 2203-2208, 'X', 2210-2211, 'X', 2213-2227; 2228-2251; 2314
A:Cross-references: GB:X06196; NID:G30055; PIDN:CAA29557.1; PID:G1335034
A:Note: the mRNA portion of the sequence corresponds to residues 2092-2157
R:Chu, M.L.; Conway, D.; Pan, T.; Baldwin, C.; Mann, K.; Deutzmann, R.; Timpl, R.
J. Biol. Chem. 263, 18601-18606, 1988
A:Title: Amino acid sequence of the triple-helical domain of human collagen type VI.
A:Reference number: A31952; MUID:89066644; PMID:3198591
A:Accession: C31952
A:Molecule type: mRNA
A:Residues: 2038-2373 <CH4>
A:Cross-references: GB:J04211; GB:M20778
A:Note: parts of this sequence were determined by protein sequencing
R:Weil, D.; Mattel, M.G.; Passage, E.; Van Cong, N.; Pribula-Conway, D.; Mann, K.; Deutz
Am. J. Hum. Genet. 42, 435-445, 1988
A:Title: Cloning and chromosomal localization of human genes encoding the three chains of
A:Reference number: A29848; MUID:88161046; PMID:3348212
A:Accession: C29848
A:Molecule type: mRNA
A:Residues: 2092-2151 <WEI>
A:Cross-references: GB:M27449; NID:G291919; PIDN:AAAS2057.1; PID:G291920
A:Note: part of this sequence was determined by protein sequencing
R:Jander, R.; Rautenberg, J.; Glanville, R.W.
Eur. J. Biochem. 133, 39-46, 1983
A:Title: Further characterization of the three polypeptide chains of bovine and human shc
A:Reference number: S26506; MUID:8309648; PMID:6852053
A:Accession: S26510
A:Molecule type: protein
```

A;Residues: 'SALAGVAGV' <JAN>
 A;Note: this sequence cannot be reliably placed and probably represents the results from
 R;Mayer, U.; Poeschl, E.; Nisch, U.; Specks, U.; Pan, T.C.; Chu, M.L.; Timpi, R.
 Eur. J. Biochem. 225, 573-580, 1994
 A;Title: Recombinant expression and properties of the Kunitz-type protease-inhibitor mod
 A;Reference number: S48709; MUID:95045506; PMID:7525281
 A;Accession: S48709
 A;Molecule type: mRNA
 A;Residues: 'NRNIFELLCAGALAA' 3102-3176 <MAY>
 A;Note: engineered sequence to allow isolation of the Kunitz-type proteinase inhibitor h
 R;Arnoux, B.; Merigau, K.; Saludivian, P.; Norris, F.; Norris, K.; Bjorn, S.; Olsen, O.
 submitted to the Brookhaven Protein Data Bank, August 1994
 A;Reference number: A52812; PDB:1XNT
 A;Contents: annotation: X-ray crystallography, 1.6 angstroms, residues 3106-3160
 A;Note: engineered sequence expressed in *Saccharomyces cerevisiae* strain mt-663
 C;Comment: Prolines and lysines at the third position of the tripeptide repeating unit
 sines are 5-hydroxylated and subsequently O-glycosylated.
 C;Comment: The fibronectin type III repeat homology domain may be released during proces
 C;Genetics:
 A;Gene: GDB:COL6A3
 A;Cross-references: GDB:119066; OMIM:120250
 A;Map position: 2G37.3-2Q37.3
 C;Complex: type VI collagen is a heterotrimer of one alpha 1(VI) chain (see PIR:CGHUA),
 sociations among trimer amino- and carboxyl-terminal domains (with disulfide bonds).
 C;Function:
 A;Description: structural component of extracellular tissue microfibrils associated with
 C;Superfamily: collagen VI
 F;1-25/Domain: signal sequence #status predicted <SIG>
 F;26-3176/Product: collagen alpha 3(VI) chain #status predicted <WAT>
 F;26-2037/Domain: amino-terminal nonhelical #status predicted <ANH>
 F;26-30-237-3176/Product: collagen alpha 3(VI) chain, splice form A9/N10(-) #status pred
 F;37-203/Domain: von Willebrand factor type A repeat homology <WV01>
 F;240-405/Domain: von Willebrand factor type A repeat homology <WV02>
 F;443-608/Domain: von Willebrand factor type A repeat homology <WV03>
 F;637-802/Domain: von Willebrand factor type A repeat homology <WV04>
 F;835-999/Domain: von Willebrand factor type A repeat homology <WV05>
 F;1027-1191/Domain: von Willebrand factor type A repeat homology <WV06>
 F;1231-1394/Domain: von Willebrand factor type A repeat homology <WV07>
 F;1434-1599/Domain: von Willebrand factor type A repeat homology <WV08>
 F;1637-1802/Domain: von Willebrand factor type A repeat homology <WV09>
 F;1836-2005/Domain: von Willebrand factor type A repeat homology <WV10>
 F;2038-2373/Region: interrupted helical
 F;2040-2042/Region: cell attachment (R-G-D) motif
 F;2136-2138/Region: cell attachment (R-G-D) motif
 F;2148-2150/Region: cell attachment (R-G-D) motif
 F;2154-2156/Region: cell attachment (R-G-D) motif
 F;2370-2372/Region: cell attachment (R-G-D) motif
 F;2374-3176/Domain: carboxyl-terminal nonhelical #status predicted <CNH>
 F;2400-2571/Domain: von Willebrand factor type A repeat homology <WV11>
 F;2617-2800/Domain: von Willebrand factor type A repeat homology <WV12>
 F;2865-2986/Region: alanine/lysine/proline/threonine/valine-rich repeats
 F;2987-3072/Region: fibronectin type III repeat homology <FN3>
 F;3111-3161/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>
 F;26/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
 F;108.116.202.251.2079.2331.2559.2677.2861.3036/Binding site: carbohydrate (Asn) (coval
 F;2087/Disulfide bonds: interchain #status predicted
 F;2100.2206.2239.2316.2319/Modified site: 4-hydroxyproline (Pro) #status experimental
 F;2103.2209.2212.2322.2337/Binding site: carbohydrate (Lys) (covalent) #status experimen
 F;2103.2209.2212.2322.2337/Modified site: 5-hydroxylysine (Lys) #status experimental
 F;3111-3161,3120-3144,3136-3157/Disulfide bonds: #status predicted

Query Match 43.6%; Score 48; DB 2; Length 3176;
 Best Local Similarity 50.0%; Pred. No. 2.2e+02;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 4 VGVFIDELGAVSDLHRA 21

Db 1948 VIHFTDGDGDLADLHRA 1965

RESULT 11
 A47486

glutathione transferase (EC 2.5.1.18) class mu, GSTM4 (version 1) - human
 C;Species: Homo sapiens (man)
 C;Date: 02-Jun-1995 #sequence_revision 02-Jun-1995 #text_change 09-Jul-2004
 C;Accession: A47486; S36782; I37438; S45685
 R;Comstock, K.E.; Johnson, K.J.; Riferbery, D.; Henner, W.D.
 J. Biol. Chem. 268, 16958-16965, 1993
 A;Title: Isolation and analysis of the gene and cDNA for a human mu class glutathione S-t
 A;Reference number: A47486; MUID:93352467; PMID:8349596
 A;Accession: A47486
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-218 <COM>
 A;Cross-references: UNIPROT:Q03013; GB:M96233
 A;Accession: B47486
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-218 <CO2>
 A;Cross-references: GB:M96234; NID:G306818; PIDN:AAA57347.1; PID:G306819
 A;Experimental source: Hela cells
 R;Ross, V.L.; Board, P.G.
 Biochem. J. 294, 373-380, 1993
 A;Title: Molecular cloning and heterologous expression of an alternatively spliced human
 A;Reference number: S36782; MUID:93384505; PMID:8373352
 A;Accession: S36782
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-218 <ROS>
 A;Cross-references: EMBL:M99422
 A;Experimental source: testis
 R;Comstock, K.E.; Widersten, M.; Hao, X.Y.; Henner, W.D.; Mannervik, B.
 Arch. Biochem. Biophys. 311, 487-495, 1994
 A;Title: A comparison of the enzymatic and physicochemical properties of human glutathior
 A;Reference number: S45684; MUID:94263230; PMID:8203914
 A;Contents: annotation
 R;Taylor, J.B.; Oliver, J.; Sherrington, R.; Pembie, S.E.
 Biochem. J. 274, 587-593, 1991
 A;Title: Structure of human glutathione S-transferase class Mu genes.
 A;Reference number: I37438; MUID:91174774; PMID:2006920
 A;Accession: I37438
 A;Status: preliminary; translated from GB/EMBL/DDEJ
 A;Molecule type: DNA
 A;Residues: 39-120 <RES>
 A;Cross-references: EMBL:X56837; NID:G31936; PIDN:CAA40167.1; PID:G31937
 C;Genetics:
 A;Gene: GDB:GSTM4
 A;Cross-references: GDB:I34191; OMIM:138333
 A;Map position: I313.3-I313.3
 A;Introns: 12/3; 38/1; 59/3; 87/1; 120/3; 152/3; 189/3
 C;Superfamily: glutathione transferase
 C;Keywords: alternative splicing; dimer; transferase

Query Match 42.7%; Score 47; DB 2; Length 218;
 Best Local Similarity 50.0%; Pred. No. 17;
 Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MGVGVFIDELGAVSDLHR 20
 Db 149 VGDKITFVDFLAYDVLHRLH 168

RESULT 12
 S32425
 glutathione transferase (EC 2.5.1.18) class mu, GSTM4 (version 2) - human
 N;Alternate names: glutathione transferase mu4
 C;Species: Homo sapiens (man)
 C;Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 09-Jul-2004
 C;Accession: S32425; S29337
 R;Zhong, S.; Spurr, N.K.; Hayes, J.D.; Wolf, C.R.
 Biochem. J. 291, 41-50, 1993
 A;Title: Deduced amino acid sequence, gene structure and chromosomal location of a novel
 A;Reference number: S32424; MUID:93228631; PMID:8471052
 A;Accession: S32425
 A;Molecule type: DNA

```

C:Genetics:
A:Map position: 1

Query Match      42.7%; Score 47; DB 2; Length 450;
Best Local Similarity 36.4%; Pred. No. 38;
Matches 8; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 1 MGDVVGFIDELEGAVSDLHRL 22
   |||::|||:|||||
Db 302 LGDAIDYKELLQRLNDLHTEL 323

RESULT 15
AG2288
GTP-binding protein [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AG2288
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AG2288
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-528 <KUR>
A:Cross-references: UNIPROT:Q8YQG9; GB:BA000019; PIDN:BAE75561.1; PID:g17132996; GSPDB:GCI
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all3862

Query Match      42.7%; Score 47; DB 2; Length 528;
Best Local Similarity 52.6%; Pred. No. 45;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 3 DWVGFIDELEGAVSDLHRA 21
   ||||| |||: |||
Db 413 DTVGFIHLPASLMDAPRA 431

Search completed: November 10, 2004, 15:55:10
Job time : 8.51601 secs

```

This Page Blank (uspio)

RL DNA Cell Biol. 15:407-414 (1996).

DR EMBL; X97286; CAA65949.1; --
 DR HSP; P02019; IOUT
 DR GO; GO:0005333; C-hemoglobin complex; IEA.
 DR GO; GO:0005344; F-oxygen transporter activity; IEA.
 DR GO; GO:0015671; P-oxygen transport; IEA.
 DR GO; GO:0006810; P-transport; IEA.
 DR InterPro; IPR002338; Alpha haem.
 DR InterPro; IPR000971; Globin.
 DR InterPro; IPR009050; Globin-like.
 DR InterPro; IPR002339; Pi-haem.
 DR Pfam; PF00042; Globin; 1.
 DR PRINTS; PR00612; ALPHAHAE.
 DR PRINTS; PR00815; PIHAEM.
 DR PROSITE; PS01033; GLOBIN; 1.
 DR Heme; Oxygen transport; Transport.
 DR KW Heme; Oxygen transport; Transport.
 DR SEQUENCE 143 AA; 15141 MW; 64D4A28956019EC3 CRC64;

Query Match 48.8%; Score 53.5; DB 2; Length 143;
 Best Local Similarity 54.5%; Pred. No. 8.5;
 Matches 12; Conservative 4; Mismatches 3; Indels 3; Gaps 1;

Qy 1 MGDVGVFIDELEG---AVSDLH 19
 Db 68 IGNAVGLIDVGLSALSDLH 89

RESULT 3
 Q83G19 PRELIMINARY; PRT; 603 AA.
 AC Q83G19;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE DNA primase (EC 2.7.7.-)
 GN Name=dnaG; OrderedLocusNames=TW288;
 OS Tropheryma whipplei (strain Twist) (Whipple's bacillus).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Micrococcales; Cellulomonadaceae; Tropheryma.
 OC NCBI_TaxID=203267;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Twist;
 RX MEDLINE=22784088; PubMed=12902375;
 RA Racult D., Ogata H., Audic S., Robert C., Suhre K., Drancourt M.,
 RA Clavierie J.-M.;
 RT "Tropheryma whipplei Twist: a human pathogenic Actinobacteria with a
 RT reduced genome.";
 RL Genome Res. 13:1800-1809(2003).
 RL EMBL; AB016851; AAO44385.1; --
 DR HSP; Q8X4D0; IDOQ.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0003896; F:DNA primase activity; IEA.
 DR GO; GO:0016740; F:nucleotidyltransferase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006304; P:DNA modification; IEA.
 DR GO; GO:0006260; P:DNA replication; IEA.
 DR GO; GO:0006269; P:DNA replication, synthesis of RNA primer; IEA.
 DR InterPro; IPR006295; DNA primase.
 DR InterPro; IPR006171; Toprim_dom.
 DR InterPro; IPR006647; Toprim primase.
 DR InterPro; IPR002694; Znf_CHC2; 1.
 DR Pfam; PF01751; Toprim; 1.
 DR ProDom; PD002276; Toprim primase; 1.
 DR ProDom; PD002988; Znf_CHC2; 1.
 DR SMART; SM00493; TOPRIN; 1.
 DR SMART; SM00400; Znf_CHCC; 1.
 DR TIGRFAMs; TIGR01391; dnaG; 1.
 DR Complete proteome; Nucleotidyltransferase; Transferase.
 DR SEQUENCE 603 AA; 67692 MW; 2E5502DA58D28072 CRC64;

Query Match 47.7%; Score 52.5; DB 2; Length 603;
 Best Local Similarity 46.2%; Pred. No. 47;
 Matches 12; Conservative 3; Mismatches 6; Indels 5; Gaps 1;

Qy 1 MGDVGVFIDELEG---AVSDLH 19
 Db 68 IGNAVGLIDVGLSALSDLH 89

RESULT 3
 Q83G19 PRELIMINARY; PRT; 603 AA.
 AC Q83G19;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE DNA primase (EC 2.7.7.-)
 GN Name=dnaG; OrderedLocusNames=TW288;
 OS Tropheryma whipplei (strain Twist) (Whipple's bacillus).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Micrococcales; Cellulomonadaceae; Tropheryma.
 OC NCBI_TaxID=203267;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Twist;
 RX MEDLINE=22784088; PubMed=12902375;
 RA Racult D., Ogata H., Audic S., Robert C., Suhre K., Drancourt M.,
 RA Clavierie J.-M.;
 RT "Tropheryma whipplei Twist: a human pathogenic Actinobacteria with a
 RT reduced genome.";
 RL Genome Res. 13:1800-1809(2003).
 RL EMBL; AB016851; AAO44385.1; --
 DR HSP; Q8X4D0; IDOQ.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0003896; F:DNA primase activity; IEA.
 DR GO; GO:0016740; F:nucleotidyltransferase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006304; P:DNA modification; IEA.
 DR GO; GO:0006260; P:DNA replication; IEA.
 DR GO; GO:0006269; P:DNA replication, synthesis of RNA primer; IEA.
 DR InterPro; IPR006295; DNA primase.
 DR InterPro; IPR006171; Toprim_dom.
 DR InterPro; IPR006647; Toprim primase.
 DR InterPro; IPR002694; Znf_CHC2; 1.
 DR Pfam; PF01751; Toprim; 1.
 DR ProDom; PD002276; Toprim primase; 1.
 DR ProDom; PD002988; Znf_CHC2; 1.
 DR TIGRFAMs; TIGR01391; dnaG; 1.
 DR Complete proteome; Nucleotidyltransferase; Transferase.
 DR SEQUENCE 603 AA; 67692 MW; 2E5502DA58D28072 CRC64;

Query Match 47.7%; Score 52.5; DB 2; Length 603;
 Best Local Similarity 46.2%; Pred. No. 47;
 Matches 12; Conservative 3; Mismatches 6; Indels 5; Gaps 1;

Qy 1 MGDVGVFIDELEG---AVSDLH 19
 Db 68 IGNAVGLIDVGLSALSDLH 89

Qy 2 GDVVGFIIDELEG---AVSDLHREAL 22
 Db 67 GDVISFVQIEIENFVDAVEQLARTL 92

RESULT 4
 Q83HP0 PRELIMINARY; PRT; 603 AA.
 AC Q83HP0;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE DNA primase (EC 2.7.7.-)
 GN Name=dnaG; OrderedLocusNames=TW484;
 OS Tropheryma whipplei (strain TW08/27) (Whipple's bacillus).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Micrococcales; Cellulomonadaceae; Tropheryma.
 OC NCBI_TaxID=218496;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=TW08/27;
 RX MEDLINE=22495039; PubMed=12606174;
 RA Bentley S.D., Maiwald M., Murphy L.D., Pallen M.J., Yeats C.A.,
 RA Dover L.G., Norbertczak H.T., Besra G.S., Quail M.A., Harris D.E.,
 RA von Herbay A., Gobie A., Rutter S., Squares R., Squares S.,
 RA Barrell B.G., Parkhill J., Rellman D.A.;
 RT "Sequencing and analysis of the genome of the Whipple's disease
 RT bacterium Tropheryma whipplei.";
 RL Lancet 361:637-644(2003).
 RL EMBL; BX251411; CAD67151.1; --
 DR HSP; Q8X4D0; IDOQ.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0003896; F:DNA primase activity; IEA.
 DR GO; GO:0016740; F:nucleotidyltransferase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006304; P:DNA modification; IEA.
 DR GO; GO:0006260; P:DNA replication; IEA.
 DR GO; GO:0006269; P:DNA replication, synthesis of RNA primer; IEA.
 DR InterPro; IPR006295; DNA primase.
 DR InterPro; IPR006171; Toprim_dom.
 DR InterPro; IPR006647; Toprim primase.
 DR InterPro; IPR006154; Toprim sub.
 DR InterPro; IPR002694; Znf_CHC2.
 DR Pfam; PF01751; Toprim; 1.
 DR ProDom; PD002276; Toprim primase; 1.
 DR ProDom; PD002988; Znf_CHC2; 1.
 DR SMART; SM00493; TOPRIN; 1.
 DR SMART; SM00400; Znf_CHCC; 1.
 DR TIGRFAMs; TIGR01391; dnaG; 1.
 DR Complete proteome; Nucleotidyltransferase; Transferase.
 DR SEQUENCE 603 AA; 67692 MW; 2E5502DA58D28072 CRC64;

Query Match 47.7%; Score 52.5; DB 2; Length 603;
 Best Local Similarity 46.2%; Pred. No. 47;
 Matches 12; Conservative 3; Mismatches 6; Indels 5; Gaps 1;

Qy 2 GDVVGFIIDELEG---AVSDLHREAL 22
 Db 67 GDVISFVQIEIENFVDAVEQLARTL 92

RESULT 5
 H5AA_SEROU STANDARD; PRT; 142 AA.
 AC Q83HP0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Hemoglobin alpha-A chain.
 GN Name=H5AA;
 OS Seriola quinqueradiata (Five-ray yellowtail).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 CC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Carangoidae;
 CC Carangidae; Seriola.
 OX NCBI_TaxID=8161;
 RN [1]_TaxID=8161;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Sakai M., Okamoto K.;
 RT "Yellowtail's mRNA for hemoglobin alpha chain A.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Involved in oxygen transport from the lung to the
 CC various peripheral tissues.
 CC -!- SUBUNIT: Heterotrimer of two alpha chains and two beta chains.
 CC -!- TISSUE SPECIFICITY: Red blood cells.
 CC -!- SIMILARITY: Belongs to the globin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AB034639; BAA86218.1; -;
 DR HSSP; P56250; 1SPG.
 DR InterPro; IPR002338; Alpha.haem.
 DR InterPro; IPR000971; Globin_like.
 DR Pfam; PF009050; Globin_1.
 DR PRINTS; PR00612; ALPHAHAM.
 DR PROSITE; PS01033; GLOBIN; 1.
 KW Erythrocyte; Heme; Oxygen transport;
 FT INIT_MET 0 By similarity.
 FT METAL 59 Iron (heme distal ligand).
 FT METAL 88 Iron (heme proximal ligand).
 SQ SEQUENCE 142 AA; 15518 MW; 059246F5463582D6 CRC64;
 Query Match 46.8%; Score 51.5; DB 1; Length 142;
 Best Local Similarity 54.5%; Pred. No. 17;
 Matches 12; Conservative 4; Mismatches 3; Indels 3; Gaps 1;
 QY 1 MGDVVGFIIDELEG---AVSDLH 19
 Db 67 VGDAVGKIDDLVGLGSLSELH 88
 RESULT 6
 QY257
 ID Q6Y257 PRELIMINARY; PRT; 143 AA.
 AC Q6Y257
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE Alpha hemoglobin A.
 OS Pagrus major (Red sea bream) (Chrysophrys major).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 CC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
 CC Sparidae; Pagrus.
 OX NCBI_TaxID=143350;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RA Chen S.L., Xu M.Y.;
 RT Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to the syntaxin/epimorphin family.
 CC Nature 425:973-977(2003).
 DR EMBL; AY190679; AAP20155.1; -;
 DR InterPro; IPR002338; Alpha.haem.
 DR InterPro; IPR000971; Globin_like.
 DR Pfam; PF00042; Globin; 1.
 DR PRINTS; PR00612; ALPHAHAM.
 DR PROSITE; PS01033; GLOBIN; 1.

KW Heme; Oxygen transport; Transport.
 SQ SEQUENCE 143 AA; 15701 MW; FDCA7EBD81968DPE CRC64;
 Query Match 46.8%; Score 51.5; DB 2; Length 143;
 Best Local Similarity 54.5%; Pred. No. 17;
 Matches 12; Conservative 3; Mismatches 4; Indels 3; Gaps 1;
 QY 1 MGDVVGFIIDELEG---VSDLH 19
 Db 68 IGDVVGKIDDLVGLGSLNLSDLH 89
 RESULT 7
 AAP20155
 ID AAP20155 PRELIMINARY; PRT; 143 AA.
 AC AAP20155;
 DT 02-MAR-2004 (TREMBLrel. 27, Created)
 DT 02-MAR-2004 (TREMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (TREMBLrel. 27, Last annotation update)
 DE Alpha hemoglobin A.
 OS Pagrus major (Red sea bream) (Chrysophrys major).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 CC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
 CC Sparidae; Pagrus.
 OX NCBI_TaxID=143350;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen; N.A.
 RA Chen S.L., Xu M.Y.;
 RT "Analysis of expressed genes in red sea bream (Chrysophrys major).";
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY190679; AAP20155.1; -;
 SQ SEQUENCE 143 AA; 15701 MW; FDCA7EBD81968DPE CRC64;
 Query Match 46.8%; Score 51.5; DB 2; Length 143;
 Best Local Similarity 54.5%; Pred. No. 17;
 Matches 12; Conservative 3; Mismatches 4; Indels 3; Gaps 1;
 QY 1 MGDVVGFIIDELEG---VSDLH 19
 Db 68 IGDVVGKIDDLVGLGSLNLSDLH 89
 RESULT 8
 Q6X9V9
 ID Q6X9V9 PRELIMINARY; PRT; 318 AA.
 AC Q6X9V9
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE Syntaxin.
 GN Name=Ror2;
 OS Hordeum vulgare var. distichum (Two-rowed barley).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 CC Triticeae; Hordeum.
 OX NCBI_TaxID=112509;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22948933; PubMed=14586469;
 RA Collins N.C., Thordal-Christensen H., Lipka V., Bau S., Kombrink E.,
 RA Qiu J.L., Huckelhoven R., Stein M., Freialdenhoven A.,
 RA Somerville S.C., Schulze-Lefert P.;
 RT "SNARE-protein-mediated disease resistance at the plant cell wall.";
 RL Nature 425:973-977(2003).
 CC -!- SIMILARITY: Belongs to the syntaxin/epimorphin family.
 DR EMBL; AY246907; AAP75622.1; -;
 DR InterPro; IPR006012; Syntaxin.
 DR InterPro; IPR006011; Syntaxin_N.
 DR InterPro; IPR010989; t-snare.
 DR InterPro; IPR000727; t-SNARE.

```

RL Nature 425:973-977(2003).
DR EMBL; AY246907; AAP75622.1; -.
SQ SEQUENCE 318 AA; 34547 MW; 908B40B242CE0C39 CRC64;

Query Match          46.8%; Score 51.5; DB 2; Length 318;
Best Local Similarity 45.8%; Pred.No.36;
Matches 11; Conservative      3; Mismatches 7; Indels 3; Gaps 1;

QY 2 GDVVGFIDELE---GAVSDLHRLAL 22
   |:|::||::|||::|||:|
Db 206 GEVLGVVAIEQRHGAVADLERSL 229

RESULT 11
Q9ABD1 PRELIMINARY; PRT; 197 AA.
ID Q9ABD1
AC Q9ABD1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT DT 01-OCT-2002 (TrEMBLrel. 25, Last annotation update)
DN Hypothetical protein CC0298.
DE OrderedLocusNames=CC0298;
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;
RA Niernan W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Neilson K.E.,
RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Neilson W.C., Newton A.S., Stephens C., Bhadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Knouri H.M., Shetty J.,
RA Berry K.J., Uterback T.R., Tran K., Wolf A.M., Vamathevan J.J.,
RA Ermolaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,
RA Fraser C.M.;
RA "Complete genome sequence of Caulobacter crescentus." ;
RT Proc Natl Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AS005703; AAK22285.1; -.
DR FICR; A87286; A87286.
DR TIGR; CC0298; -.
DR InterPro; IPR004360; Gly_bleo_diox.
DR Pfam; PF00903; Glyoxalase; 1_
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 197 AA; 21365 MW; 68ED283154210409 CRC64;

Query Match          46.4%; Score 51; DB 2; Length 197;
Best Local Similarity 61.1%; Pred.No.27;
Matches 11; Conservative      2; Mismatches 5; Indels 0; Gaps 0;

QY 5 VGFIDELEGAVSDLHRLAL 22
   |:|::|||::|||:|
Db 145 IGPHLETEEAVSDLRHM 162

RESULT 12
Q8Q0A2 PRELIMINARY; PRT; 425 AA.
ID Q8Q0A2
AC Q8Q0A2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Conserved protein.
DE OrderedLocusNames=MW0235;
GN Methanosarcina mazei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2205;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Goe1 / Go1 / ATCC BAA-199 / DSM 3647 / OCM 88;

```

```

RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wieser A., Baumer S., Jacobi C.,
RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,
RA Brattacharya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of Methanosarcina mazei: evidence for lateral gene
RL transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
DR EMBL; AE013248; AAM29931.1; -.
KW Complete proteome.
SQ SEQUENCE 425 AA; 47957 MW; 00B0CBF7B7C3DB0D CRC64;

Query Match 46.4%; Score 51; DB 2; Length 425;
Best Local Similarity 50.8%; Pred. No. 57;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 GDVVVGFDLEAGVSDIHLRA 21
Db 45 GDVSGLFKEIAAISLRES 64

RESULT 13
Q73Q16 PRELIMINARY; PRT; 646 AA.
AC Q73Q16
DT 03-JUN-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DE Chapterone protein DnaK.
GN Names-dnaK; OrderedLocusNames-TDS0628;
OS Treponema denticola.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=158;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35405 / DSM 14222;
RA Seshadri R., Myers G.S.A., Tettelin H., Eisen J.A., Heidelberg J.F.,
RA Dodson R.J., Davidson T.M., DeBoy R.T., Fouts D.E., Haft D.H.,
RA Selengut J., Ren Q., Brinkac L.M., Madupu R., Kolonay J.F.,
RA Daugherty S.C., Shetty J., Shvartsbeyn A., Gebregeorgis E.,
RA Tsegaye G., Malek J.A., Ayodeji B., Shatsman S., McLeod M.P.,
RA Shatsman S., McLeod M.P., Smajls D., Howell J.K., Pal S., Amin A.,
RA Vashisth P., McNeill T.Z., Xiang Q., Sodergren E., Baca E.,
RA Weinstein G.M., Norris S.J., Fraser C.M., Paulsen I.T.;
RT "Comparison of the genome of the oral pathogen Treponema denticola
RT with other spirochete genomes.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:5646-5651(2004).
DR EMBL; AE017248; AAS11123.1; -.
DR TIGR; TDS0628; -.
SQ SEQUENCE 646 AA; 69759 MW; 2A61E0B4C74F06C8 CRC64;

Query Match 46.4%; Score 51; DB 2; Length 646;
Best Local Similarity 45.8%; Pred. No. 84;
Matches 11; Conservative 6; Mismatches 5; Indels 2; Gaps 1;

QY 1 MGDVVVGFD--ELEGAVSDIHLRAL 22
Db 547 MGDKIGAADKQKIEAAIADLRQAL 570

RESULT 15
Q9LLD4 PRELIMINARY; PRT; 1824 AA.
ID Q9LLD4
AC Q9LLD4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Pfl.
OS Lycopersicon pimpinellifolium (Current tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4084;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21948620; PubMed=11952131;
RA Chang J.H., Tai Y.S., Bernal A.J., Lavelle D.T., Staskawicz B.J.,
RA Michelmore R.W.;
RT "Functional analyses of the Pto resistance gene family in tomato and
RT the identification of a minor resistance determinant in a susceptible
RT haplotype.";
RL Mol. Plant Microbe Interact. 15:281-291(2002).
DR EMBL; AF220602; AAF76308.1; -.
DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:0006915; P:apoptosis; IEA.
DR GO; GO:0042829; P:defense response to pathogen; IEA.
DR InterPro; IPR000767; Disease resist.
DR InterPro; IPR001916; Glyco_hydro_22.
DR InterPro; IPR002182; NB-ARC.
DR Pfam; PF00931; NB-ARC; 1.
DR PRINTS; PR00364; DISEASERSIST.
DR PROSITE; PS00128; LACTALBUMIN LYSOZYME; UNKNOWN 1.
SQ SEQUENCE 1824 AA; 209573 MW; 1F14E7D662DE6AC6 CRC64;

Query Match 45.9%; Score 50.5; DB 2; Length 1824;

```

```

ID AAS11123 PRELIMINARY; PRT; 646 AA.
AC AAS11123;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 04-MAY-2004 (TREMBlrel. 27, Last annotation update)
DE Chapterone protein DnaK.
GN DnaK OR TDS0628.
OS Treponema denticola.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=158;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35405 / DSM 14222;
RA Seshadri R., Myers G.S.A., Tettelin H., Eisen J.A., Heidelberg J.F.,
RA Dodson R.J., Davidson T.M., DeBoy R.T., Fouts D.E., Haft D.H.,
RA Selengut J., Ren Q., Brinkac L.M., Madupu R., Kolonay J.F.,
RA Daugherty S.C., Shetty J., Shvartsbeyn A., Gebregeorgis E.,
RA Tsegaye G., Malek J.A., Ayodeji B., Shatsman S., McLeod M.P.,
RA Smajls D., Howell J.K., Pal S., Amin A., Vashisth P., McNeill T.Z.,
RA Xiang Q., Sodergren E., Baca E., Weinstein G.M., Norris S.J.,
RA Fraser C.M., Paulsen I.T.;
RT "Comparison of the genome of the oral pathogen Treponema denticola
RT with other spirochete genomes.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:5646-5651(2004).
DR EMBL; AE017248; AAS11123.1; -.
DR TIGR; TDS0628; -.
SQ SEQUENCE 646 AA; 69759 MW; 2A61E0B4C74F06C8 CRC64;

Query Match 46.4%; Score 51; DB 2; Length 646;
Best Local Similarity 45.8%; Pred. No. 84;
Matches 11; Conservative 6; Mismatches 5; Indels 2; Gaps 1;

QY 1 MGDVVVGFD--ELEGAVSDIHLRAL 22
Db 547 MGDKIGAADKQKIEAAIADLRQAL 570

RESULT 15
Q9LLD4 PRELIMINARY; PRT; 1824 AA.
ID Q9LLD4
AC Q9LLD4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Pfl.
OS Lycopersicon pimpinellifolium (Current tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4084;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21948620; PubMed=11952131;
RA Chang J.H., Tai Y.S., Bernal A.J., Lavelle D.T., Staskawicz B.J.,
RA Michelmore R.W.;
RT "Functional analyses of the Pto resistance gene family in tomato and
RT the identification of a minor resistance determinant in a susceptible
RT haplotype.";
RL Mol. Plant Microbe Interact. 15:281-291(2002).
DR EMBL; AF220602; AAF76308.1; -.
DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:0006915; P:apoptosis; IEA.
DR GO; GO:0042829; P:defense response to pathogen; IEA.
DR InterPro; IPR000767; Disease resist.
DR InterPro; IPR001916; Glyco_hydro_22.
DR InterPro; IPR002182; NB-ARC.
DR Pfam; PF00931; NB-ARC; 1.
DR PRINTS; PR00364; DISEASERSIST.
DR PROSITE; PS00128; LACTALBUMIN LYSOZYME; UNKNOWN 1.
SQ SEQUENCE 1824 AA; 209573 MW; 1F14E7D662DE6AC6 CRC64;

Query Match 45.9%; Score 50.5; DB 2; Length 1824;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:55:47 ; Search time 6.9395 Seconds
(without alignments)
143.349 Million cell updates/sec

Title: US-10-092-750-228

Perfect score: 86

Sequence: 1 LRHWGLQFNTRFGV 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/1aa/5A-COMB.pep:*
- 2: /cgn2_6/prodata/1/1aa/5B-COMB.pep:*
- 3: /cgn2_6/prodata/1/1aa/5A-COMB.pep:*
- 4: /cgn2_6/prodata/1/1aa/5B-COMB.pep:*
- 5: /cgn2_6/prodata/1/1aa/PCTUS-COMB.pep:*
- 6: /cgn2_6/prodata/1/1aa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 62 | 72.1 | 51 | 4 | US-09-621-976-7555 |
| 2 | 44 | 51.2 | 149 | 4 | US-09-543-681A-8305 |
| 3 | 44 | 51.2 | 263 | 4 | US-09-173-300-20 |
| 4 | 43 | 50.0 | 311 | 4 | US-09-252-991A-22062 |
| 5 | 41 | 47.7 | 265 | 4 | US-09-583-110-4596 |
| 6 | 40 | 46.5 | 159 | 4 | US-09-270-767-42163 |
| 7 | 40 | 46.5 | 322 | 4 | US-09-540-236-2700 |
| 8 | 40 | 46.5 | 674 | 3 | US-09-199-637A-51 |
| 9 | 40 | 46.5 | 674 | 4 | US-09-252-991A-26476 |
| 10 | 40 | 46.5 | 737 | 4 | US-09-248-796A-14319 |
| 11 | 40 | 46.5 | 1463 | 1 | US-08-157-005-3 |
| 12 | 40 | 46.5 | 1463 | 3 | US-08-747-863-3 |
| 13 | 40 | 46.5 | 1463 | 4 | US-09-565-864-3 |
| 14 | 39 | 45.3 | 277 | 4 | US-09-543-681A-4527 |
| 15 | 39 | 45.3 | 321 | 4 | US-09-489-039A-10546 |
| 16 | 39 | 45.3 | 326 | 4 | US-09-543-681A-6340 |
| 17 | 39 | 45.3 | 526 | 4 | US-09-489-039A-14342 |
| 18 | 38.5 | 44.8 | 160 | 4 | US-09-270-767-49066 |
| 19 | 38.5 | 44.8 | 161 | 4 | US-09-270-767-33849 |
| 20 | 38 | 44.2 | 10 | 4 | US-09-646-423-2 |
| 21 | 38 | 44.2 | 10 | 4 | US-09-428-082B-1094 |
| 22 | 38 | 44.2 | 105 | 3 | US-08-936-165A-439 |
| 23 | 38 | 44.2 | 163 | 4 | US-09-621-976-6321 |
| 24 | 38 | 44.2 | 449 | 4 | US-09-489-039A-8009 |
| 25 | 37.5 | 43.6 | 142 | 5 | PCT-US95-06266-134 |
| 26 | 37.5 | 43.6 | 175 | 5 | PCT-US95-06266-147 |
| 27 | 37 | 43.0 | 61 | 4 | US-09-248-796A-27758 |

ALIGNMENTS

RESULT 1

US-09-621-976-7555
; Sequence 7555, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTS and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 7555
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-7555

Query Match 72.1%; Score 62; DB 4; Length 51;
Best Local Similarity 76.9%; Pred. No. 0.0014;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

| | | | |
|----|----|---------------|----|
| Qy | 2 | LRHWGLQFNTRFG | 14 |
| Db | 27 | LQHWGLHFNTRFG | 39 |

RESULT 2

US-09-543-681A-8305
; Sequence 8305, Application US/09543681A
; Patent No. 8605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 8305
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-8305

Query Match 51.2%; Score 44; DB 4; Length 149;

Sequence 4423, Ap
Sequence 3180, A
Sequence 48397, A
Sequence 3, Appli
Sequence 7987, Ap
Sequence 20713, A
Sequence 20944, A
Sequence 26, Appl
Sequence 10, Appl
Sequence 2, Appli
Sequence 6, Appli
Sequence 4, Appli
Sequence 294, App
Sequence 59, Appl
Sequence 20375, A
Sequence 29, Appl
Sequence 71, Appl
Sequence 71, Appl

28 37 43.0 75 4 US-09-134-000C-4423
29 37 43.0 144 4 US-09-270-767-33180
30 37 43.0 144 4 US-09-270-767-48397
31 37 43.0 254 3 US-09-362-831-3
32 37 43.0 344 4 US-09-328-352-7987
33 37 43.0 390 4 US-09-252-991A-20713
34 37 43.0 396 4 US-09-248-796A-20944
35 37 43.0 512 3 US-09-134-078-26
36 37 43.0 522 4 US-09-591-694-10
37 37 43.0 575 1 US-07-683-957B-2
38 37 43.0 861 4 US-09-873-737A-6
39 37 43.0 862 4 US-09-170-496D-294
40 37 43.0 1284 4 US-09-364-425B-59
41 37 43.0 1284 4 US-09-248-796A-20375
42 36.5 42.4 571 4
43 36 41.9 23 2 US-08-993-581B-29
44 36 41.9 42 2 US-08-488-161-71
45 36 41.9 42 3 US-09-273-685-71

```
Best Local Similarity 50.0%; Pred. No. 4.8;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 TLRHWGLQFNTRG 14
Db 9 TLQWAFSKGRIG 22

RESULT 3
US-09-173-300-20
; Sequence 20, Application US/09173300
; Patent No. 6451581
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Hitz, William D.
; APPLICANT: Kinney, Anthony J.
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Rafalski, J. Antoni
; TITLE OF INVENTION: PLANT BRANCHED CHAIN AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: BB-1126
; CURRENT APPLICATION NUMBER: US/09/173,300
; CURRENT FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: 60/063,423
; EARLIER FILING DATE: 1997 October 28
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 20
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-09-173-300-20

Query Match 51.2%; Score 44; DB 4; Length 363;
Best Local Similarity 63.6%; Pred. No. 13;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 LRLHWGLQFNTR 12
Db 269 LRHWGLQVSR 279

RESULT 4
US-09-252-991A-22062
; Sequence 22062, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22062
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22062

Query Match 50.0%; Score 43; DB 4; Length 311;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LRHWGLQ 8
Db 148 LRHWGLQ 154

RESULT 5
```

```
US-09-583-110-4596
; Sequence 4596, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4596
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4596

Query Match 47.7%; Score 41; DB 4; Length 265;
Best Local Similarity 58.3%; Pred. No. 29;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 LRHWGLQFNTRF 13
Db 231 LRHWGLQFNTRF 242

RESULT 6
US-09-270-767-42163
; Sequence 42163, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42163
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-42163

Query Match 46.5%; Score 40; DB 4; Length 159;
Best Local Similarity 61.5%; Pred. No. 24;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 TLRHWGLQFNTRF 13
Db 137 TVRIWPLQFFVR 149

RESULT 7
US-09-540-236-2700
; Sequence 2700, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARH
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
```

; SEQ ID NO 2700
; LENGTH: 322
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-2700

Query Match 46.5%; Score 40; DB 4; Length 322;
Best Local Similarity 35.7%; Pred. No. 53;
Matches 5; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 TLRHWGLQFNTRFG 14
Db 228 SINEWMSFQSRFG 241

RESULT 8

US-09-199-637A-51
; Sequence 51, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 674
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-51

Query Match 46.5%; Score 40; DB 3; Length 674;
Best Local Similarity 53.3%; Pred. No. 1.2e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TLRHWGLQFNTRFGV 15
Db 54 TGRHWLAANSRAGV 68

RESULT 9

US-09-252-991A-26476
; Sequence 26476, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26476
; LENGTH: 674
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26476

Query Match 46.5%; Score 40; DB 4; Length 674;
Best Local Similarity 53.3%; Pred. No. 1.2e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TLRHWGLQFNTRFGV 15
Db 54 TGRHWLAANSRAGV 68

RESULT 10

US-09-248-796A-14319
; Sequence 14319, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 14319
; LENGTH: 737
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (60)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-248-796A-14319

Query Match 46.5%; Score 40; DB 4; Length 737;
Best Local Similarity 40.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 TLRHWGLQFNTRFGV 15
Db 355 TIOHWNLNKTEFGI 369

RESULT 11

US-08-157-005-3
; Sequence 3, Application US/08157005
; Patent No. 5620691
; GENERAL INFORMATION:
; APPLICANT: Wensvoort, Gert
; APPLICANT: Terpstra, Catharinus
; APPLICANT: Pol, Johannes M
; APPLICANT: Moorman, Robertus J
; TITLE OF INVENTION: CAUSITIVE AGENT OF THE MYSTERY SWINE DISEASE,
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/157,005
; FILING DATE: 26-NOV-1993

```

; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 91201398.4
; FILING DATE: 06-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 92200781.0
; FILING DATE: 18-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/NL92/00096
; FILING DATE: 05-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Moran, Thomas F
; REGISTRATION NUMBER: 16,579
; REFERENCE/DOCKET NUMBER: 44819
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 315-1931
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1463 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-157-005-3

Query Match 46.5%; Score 40; DB 1; Length 1463;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 RHMGLOFNTRFG 14
||| |||||
Db 156 RHKGRLLNTRFG 167

RESULT 12
US-08-747-863-3
; Sequence 3, Application US/08747863
; Patent No. 6197310
; GENERAL INFORMATION:
; APPLICANT: Wensvoort, Gert
; APPLICANT: Terpstra, Catharinus
; APPLICANT: Pol, Johannes M
; APPLICANT: Moorman, Robertus J
; APPLICANT: Meulenbergh, Johanna J
; TITLE OF INVENTION: CAUSITIVE AGENT OF THE MYSTERY SWINE DISEASE,
; TITLE OF INVENTION: VACCINE COMPOSITIONS AND DIAGNOSTIC KITS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Trask, Britt & Rossa
; STREET: 525 South 300 East
; CITY: Salt Lake City
; STATE: Utah
; COUNTRY: USA
; ZIP: 84111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: WINDOWS95
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/747,863
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/157,005
; FILING DATE: 26-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 91201398.4
; FILING DATE: 06-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 92200781.0
; FILING DATE: 18-MAR-1992

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/NL92/00096
; FILING DATE: 05-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Turner, Allen C
; REGISTRATION NUMBER: 33041
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (801) 532-1922
; TELEFAX: (801) 531-9168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1463 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-747-863-3

Query Match 46.5%; Score 40; DB 3; Length 1463;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 RHMGLOFNTRFG 14
||| |||||
Db 156 RHKGRLLNTRFG 167

RESULT 13
US-09-565-864-3
; Sequence 3, Application US/09565864
; Patent No. 6455245
; GENERAL INFORMATION:
; APPLICANT: Wensvoort, Gert
; APPLICANT: Terpstra, Catharinus
; APPLICANT: Pol, Johannes M
; APPLICANT: Moorman, Robertus J
; APPLICANT: Meulenbergh, Johanna J
; TITLE OF INVENTION: CAUSITIVE AGENT OF THE MYSTERY SWINE DISEASE,
; TITLE OF INVENTION: VACCINE COMPOSITIONS AND DIAGNOSTIC KITS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/565,864
; FILING DATE: 05-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/157,005
; FILING DATE: <Unknown>
; APPLICATION NUMBER: EP 91201398.4
; FILING DATE: 06-JUN-1991
; APPLICATION NUMBER: EP 92200781.0
; FILING DATE: 18-MAR-1992
; APPLICATION NUMBER: PCT/NL92/00096
; FILING DATE: 05-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Moran, Thomas F
; REGISTRATION NUMBER: 16,579
; REFERENCE/DOCKET NUMBER: 44819
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 315-1931
; TELEX: 422523 COOP UI

```


; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1463 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-565-864-3

Query Match 46.5%; Score 40; DB 4; Length 1463;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 RHWGLOFNTRFG 14
Db 156 RHKGRLLNTRFG 167

RESULT 14
US-09-543-681A-4527
; Sequence 4527 Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4527
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4527

Query Match 45.3%; Score 39; DB 4; Length 277;
Best Local Similarity 72.7%; Pred. No. 66;
Matches 8; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 2 LRHWGLQFNTR 12
Db 177 LSHWG--FNTR 185

RESULT 15
US-09-489-039A-10546
; Sequence 10546, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10546
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10546

Query Match 45.3%; Score 39; DB 4; Length 321;
Best Local Similarity 63.6%; Pred. No. 77;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 WGLQFNTRFG 15

Db 298 WRQOFKARFGV 308

Search completed: November 10, 2004, 15:57:25
Job time : 7.9395 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 02:43:24 ; Search time 36.2456 Seconds
(without alignments)
146.426 Million cell updates/sec

Title: US-10-092-750-228

Perfect score: 86

Sequence: 1 TLRHWGLQFNTRFGV 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1568699 seqs, 353819137 residues

Total number of hits satisfying chosen parameters: 1568699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

```
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09D_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--|
| 1 | 86 | 100.0 | 15 | 14 | US-10-092-750-228 |
| 2 | 54 | 62.8 | 113 | 14 | Sequence 228, App |
| 3 | 47 | 54.7 | 684 | 15 | Sequence 5549, App |
| 4 | 44 | 51.2 | 71 | 17 | US-10-282-122A-48915 |
| 5 | 44 | 51.2 | 94 | 15 | Sequence 48915, A |
| 6 | 44 | 51.2 | 363 | 13 | US-10-425-115-187832 |
| 7 | 43 | 50.0 | 26 | 10 | Sequence 187832, Sequence 267075, Sequence 20, Appl |
| 8 | 43 | 50.0 | 26 | 14 | US-10-027-450-20 |
| 9 | 43 | 50.0 | 26 | 14 | US-09-962-756-410 |
| 10 | 43 | 50.0 | 26 | 15 | US-10-283-471-410 |
| 11 | 43 | 50.0 | 85 | 17 | US-10-253-493-410 |
| 12 | 43 | 50.0 | 223 | 17 | US-10-425-115-201421 |
| 13 | 43 | 50.0 | 308 | 14 | Sequence 201421, Sequence 357198, Sequence 8569, App |
| | | | | | Sequence 14693, A |

| | | | | | | |
|----|----|------|------|----|----------------------|--------------------|
| 14 | 42 | 48.8 | 96 | 16 | US-10-437-963-136940 | Sequence 136940, |
| 15 | 42 | 48.8 | 100 | 10 | US-09-974-879-279 | Sequence 279, App |
| 16 | 42 | 48.8 | 100 | 10 | US-09-305-736-267 | Sequence 267, App |
| 17 | 42 | 48.8 | 100 | 10 | US-09-818-683-267 | Sequence 267, App |
| 18 | 42 | 48.8 | 100 | 11 | US-09-818-683-267 | Sequence 267, App |
| 19 | 42 | 48.8 | 100 | 15 | US-10-621-401-279 | Sequence 279, App |
| 20 | 42 | 48.8 | 380 | 16 | US-10-437-963-136940 | Sequence 136940, |
| 21 | 42 | 48.8 | 431 | 15 | US-10-425-114-58855 | Sequence 58855, A |
| 22 | 42 | 48.8 | 493 | 15 | US-10-425-114-57496 | Sequence 57496, A |
| 23 | 42 | 48.8 | 493 | 17 | US-10-425-115-358150 | Sequence 358150, |
| 24 | 42 | 48.8 | 650 | 14 | US-10-369-493-5480 | Sequence 5480, App |
| 25 | 42 | 48.8 | 689 | 17 | US-10-425-115-358151 | Sequence 358151, |
| 26 | 42 | 48.8 | 792 | 15 | US-10-425-114-70483 | Sequence 70483, A |
| 27 | 42 | 48.8 | 847 | 16 | US-10-437-963-161149 | Sequence 161149, |
| 28 | 42 | 48.8 | 1137 | 14 | US-10-259-165-330 | Sequence 330, App |
| 29 | 41 | 47.7 | 70 | 9 | US-09-908-711-81 | Sequence 81, Appl |
| 30 | 41 | 47.7 | 70 | 10 | US-09-764-891-3062 | Sequence 3062, App |
| 31 | 41 | 47.7 | 265 | 9 | US-09-815-242-13270 | Sequence 13270, A |
| 32 | 41 | 47.7 | 265 | 16 | US-10-474-776-314 | Sequence 314, App |
| 33 | 41 | 47.7 | 344 | 14 | US-10-369-493-13507 | Sequence 13507, A |
| 34 | 41 | 47.7 | 346 | 15 | US-10-282-122A-62705 | Sequence 62705, A |
| 35 | 41 | 47.7 | 346 | 15 | US-10-282-122A-64833 | Sequence 64833, A |
| 36 | 41 | 47.7 | 440 | 15 | US-10-424-599-269318 | Sequence 269318, |
| 37 | 41 | 47.7 | 619 | 15 | US-10-037-417-34 | Sequence 34, Appl |
| 38 | 41 | 47.7 | 1458 | 14 | US-10-203-224-3 | Sequence 3, Appl |
| 39 | 40 | 46.5 | 45 | 9 | US-09-864-761-40200 | Sequence 40200, A |
| 40 | 40 | 46.5 | 99 | 16 | US-10-437-963-132470 | Sequence 132470, |
| 41 | 40 | 46.5 | 103 | 15 | US-10-424-599-143368 | Sequence 143368, |
| 42 | 40 | 46.5 | 104 | 15 | US-10-108-260A-3937 | Sequence 3937, App |
| 43 | 40 | 46.5 | 250 | 14 | US-10-156-761-13891 | Sequence 13891, A |
| 44 | 40 | 46.5 | 674 | 10 | US-09-975-719-51 | Sequence 51, Appl |
| 45 | 40 | 46.5 | 1463 | 14 | US-10-226-065-3 | Sequence 3, Appl |

ALIGNMENTS

RESULT 1
US-10-092-750-228
; Sequence 228, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002 US/10/092,750
; CURRENT APPLICATION NUMBER: 2002-03-07
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 228
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-228

Query Match 100.0%; Score 86; DB 14; Length 15;
Best Local Similarity 100.0%; Pred No. 9.5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TLRHWGLQFNTRFGV 15
| | | | | | | | | | | | | | | |
Db 1 TLRHWGLQFNTRFGV 15
RESULT 2
US-10-106-698-5549
; Sequence 5549, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:

QY 2 LRHWGLQFNTRF 13
| | | | | | | |
Db 16 LRHWGLQFNTRF 27

RESULT 6
US-10-027-450-20
; Sequence 20, Application US/10027450
; Publication No. US20020102715A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Hitz, William D.
; APPLICANT: Kinney, Anthony J.
; APPLICANT: Caboon, Rebecca E.
; APPLICANT: Rafalski, J. Antoni
; TITLE OF INVENTION: PLANT BRANCHED CHAIN AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: BB-1126
; CURRENT APPLICATION NUMBER: US/10/027,450
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/063,423
; PRIOR FILING DATE: 1997 October 28
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 20
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-027-450-20

Query Match 51.2%; Score 44; DB 13; Length 363;
Best Local Similarity 63.6%; Pred. No. 87;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 LRHWGLQFNTRF 12
| | | | | | | |
Db 269 LRHWGLQFNTRF 279

RESULT 7
US-09-962-756-410
; Sequence 410, Application US/09962756
; Publication No. US20030195147A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA
; APPLICANT: BRISSETTE, RENEE
; APPLICANT: BLUME, ARTHUR J.
; APPLICANT: SCHAEFFER, LAUGE
; APPLICANT: BRANDT, JAKOB
; APPLICANT: GOLDSTEIN, NEIL I.
; APPLICANT: SPETZLER, JANE
; APPLICANT: OSTERGAARD, SOREN
; APPLICANT: HANSEN, PER HERTZ
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4051U51
; CURRENT APPLICATION NUMBER: US/09/962,756
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 410
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-962-756-410

Query Match 50.0%; Score 43; DB 10; Length 26;
Best Local Similarity 63.6%; Pred. No. 9.9;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 3 RHWGLQFNTRF 13
| | | | | | | |
Db 1 RNNWLNQFNENF 11

RESULT 8
US-10-253-471-410
; Sequence 410, Application US/10253471
; Publication No. US20030236190A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4057
; CURRENT APPLICATION NUMBER: US/10/253,471
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 410
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-253-471-410

Query Match 50.0%; Score 43; DB 14; Length 26;
Best Local Similarity 63.6%; Pred. No. 9.9;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 RHWGLQFNTRF 13
| | | | | | | |
Db 1 RNNWLNQFNENF 11

RESULT 9
US-10-253-493-410
; Sequence 410, Application US/10253493
; Publication No. US20040023887A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4056
; CURRENT APPLICATION NUMBER: US/10/253,493
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 410
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-253-493-410

Query Match 50.0%; Score 43; DB 15; Length 26;
Best Local Similarity 63.6%; Pred. No. 9.9;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
QY 3 RHWGLOFNTRF 13
Db 1 RNNWLOFNENF 11

RESULT 10
US-10-425-115-201421
; Sequence 201421, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 201421
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_115280C.1.pbp
US-10-425-115-201421

Query Match 50.0%; Score 43; DB 17; Length 85;
Best Local Similarity 77.8%; Pred. No. 31;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LRHWGLQFN 10
Db 77 LRKWLGLFN 85

RESULT 11
US-10-425-115-357198
; Sequence 357198, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 357198
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_88934C.1.pbp
US-10-425-115-357198

Query Match 50.0%; Score 43; DB 17; Length 223;
Best Local Similarity 66.7%; Pred. No. 79;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LRHWGLQFN 10
Db 84 LRHWGLQFN 92

RESULT 12
US-10-156-761-8569
; Sequence 8569, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8569
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8569

Query Match 50.0%; Score 43; DB 14; Length 308;
Best Local Similarity 54.5%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 WGLQFNTRFGV 15
Db 140 WGFMYGTRFGL 150

RESULT 13
US-10-156-761-14693
; Sequence 14693, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14693
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14693

Query Match 50.0%; Score 43; DB 14; Length 323;
Best Local Similarity 88.9%; Pred. No. 1.1e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 GLQFNTRFG 14
Db 241 GLPFNTRFG 249

RESULT 14
US-10-437-963-136940
; Sequence 136940, Application US/10437963
; Publication No. US20040123343A1
```

GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 136940
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_38471C.1.pgp
US-10-437-963-136940

Query Match 48.8%; Score 42; DB 15; Length 96;
Best Local Similarity 42.9%; Pred. No. 50;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 LRHWGLQFNTRFGV 15
||| :|||
Db 28 LHHWVQKHEFTI 41

RESULT 15
US-09-974-879-279
; Sequence 279, Application US/09974879
; Publication No. US20030028003A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 125 Human Secreted Proteins
; FILE REFERENCE: P2020P2
; CURRENT APPLICATION NUMBER: US/09/974,879
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/239,893
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 09/818,683
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 09/305,736
; PRIOR FILING DATE: 1999-05-05
; PRIOR APPLICATION NUMBER: PCT/US98/23435
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: US 60/064,911
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,912
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,983
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,900
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,988
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,987
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,908
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,984
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,985
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/066,094
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,100
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,089

; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,095
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,090
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 611
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 279
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-974-879-279
Query Match 48.8%; Score 42; DB 10; Length 100;
Best Local Similarity 50.0%; Pred. No. 52;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 1 TLRWGLQFNTRFG 14
:||| :|||
Db 58 SLRSWGAKLGLRFG 71
Search completed: November 11, 2004, 07:41:47
Job time : 36.2456 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 12:29:32 ; Search time 5.79275 Seconds
(without alignments)
431.857 Million cell updates/sec

Title: US-10-092-750-31

Perfect score: 143

Sequence: 1 MRDLPGHYETLKLPLVGHKLKTIADHR 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:**

1: PIR1:**

2: PIR2:**

3: PIR3:**

4: PIR4:**

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 134 | 93.7 | 735 | 2 A59434 | KIAA1501 protein [|
| 2 | 92 | 64.3 | 1944 | 2 A59438 | KIAA1424 protein [|
| 3 | 67 | 46.9 | 733 | 2 S44876 | ZC21.4 protein - C |
| 4 | 65 | 45.5 | 1261 | 2 E59430 | PTPL1-associated R |
| 5 | 64 | 44.8 | 634 | 2 T27959 | hypothetical prote |
| 6 | 64 | 44.8 | 863 | 2 T27958 | hypothetical prote |
| 7 | 61 | 42.7 | 334 | 2 S29128 | N-chimerin - rat |
| 8 | 58 | 40.6 | 2167 | 2 S50658 | bud emergence prot |
| 9 | 56 | 39.2 | 1445 | 2 A59437 | KIAA1204 protein [|
| 10 | 55 | 38.5 | 295 | 2 B53764 | beta-chimerin - ra |
| 11 | 55 | 38.5 | 443 | 2 B53764 | beta2-chimerin, ce |
| 12 | 55 | 38.5 | 466 | 2 A53764 | hypothetical prote |
| 13 | 55 | 38.5 | 837 | 2 T19825 | hypothetical prote |
| 14 | 55 | 38.5 | 862 | 2 T46289 | hypothetical prote |
| 15 | 55 | 38.5 | 903 | 2 T00705 | N-chimerin homolog |
| 16 | 55 | 38.5 | 1003 | 2 T34066 | hypothetical prote |
| 17 | 54 | 37.8 | 299 | 2 A43953 | N-chimerin - commo |
| 18 | 54 | 37.8 | 299 | 2 S08242 | N-chimerin - human |
| 19 | 53 | 37.1 | 643 | 2 B59436 | Rho GTPase activat |
| 20 | 51 | 35.7 | 491 | 2 S74613 | hypothetical prote |
| 21 | 51 | 35.7 | 512 | 2 AF3628 | nitrate reductase |
| 22 | 50 | 35.0 | 666 | 2 S29349 | hypothetical prote |
| 23 | 49 | 34.3 | 281 | 2 T22291 | hypothetical prote |
| 24 | 49 | 34.3 | 304 | 2 D70607 | hypothetical prote |
| 25 | 49 | 34.3 | 426 | 2 D73367 | conserved hypotnet |
| 26 | 49 | 34.3 | 574 | 2 T29005 | hypothetical prote |
| 27 | 48 | 33.6 | 719 | 2 B52043 | probable glutamate |
| 28 | 48 | 33.6 | 723 | 2 B38749 | 3-phosphatidylinos |
| 29 | 48 | 33.6 | 728 | 2 T01200 | probable glutamate |

30 48 33.6 1165 2 D59433 C. elegans protein
31 48 33.6 2437 2 S42612 transmembrane prot
32 47.5 33.2 1098 2 S36868 alpha,alpha-trehal
33 326 2 AE1169 hypothetical prote
34 47 32.9 970 2 D59435 Gem-interacting pr
35 47 32.9 1358 2 B86241 hypothetical prote
36 46.5 32.5 933 2 S19427 probable membrane
37 46 32.2 290 2 T40031 hypothetical prote
38 46 32.2 718 2 T22329 hypothetical prote
39 46 32.2 764 2 JC5643 thyroid stimulat in
40 46 32.2 819 2 AH0056 aspartate kinase (
41 46 32.2 859 2 A49307 98K GTPase-activat
42 46 32.2 915 2 B59433 chromosome 5 GAP-1
43 46 32.2 1289 2 T31344 GP80 precursor - s
44 45 31.5 150 2 D72396 conserved hypotnet
45 45 31.5 331 2 S33223 transcription fact

ALIGNMENTS

RESULT 1

A59434
KIAA1501 protein [imported] - human
C/Species: Homo sapiens (man)
C/Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 09-Jul-2004
C/Accession: A59434; B59434
R/Nagase, T.; Kikuno, R.; Ishikawa, K.; Hirose, M.; Ohara, O.
mRNA Res. 7, 143-150, 2000
A/Title: Prediction of the coding sequences of unidentified human genes. XVII. The complete
A/Reference number: A59434
A/Accession: A59434
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-735 <NAG>
A/Cross-references: UNIPROT:Q9P227; GB:BAA96025; PID:g7959263; PIDN:BAA96025.1
R/Ohara, O.; Nagase, T.; Kikuno, R.
submitted to Genbank, April 2000
A/Reference number: B59434
A/Accession: B59434
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-735 <OHA>
A/Cross-references: GB:BAA96025; PID:g7959263; PIDN:BAA96025.1

Query Match 93.7%; Score 134; DB 2; Length 735;
Best Local Similarity 96.0%; Pred. No. 1.7e-11;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRDLPGHYETLKLPLVGHKLKTIADH 25
:|||||
Db 612 IRDLPGHYETLKLPLVGHKLKTIADH 636

RESULT 2

A59438
KIAA1424 protein [imported] - human
C/Species: Homo sapiens (man)
C/Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 09-Jul-2004
C/Accession: A59438; B59438
R/Nagase, T.; Kikuno, R.; Ishikawa, K.I.; Hirose, M.; Ohara, O.
DNA Res. 7, 65-73, 2000
A/Title: Prediction of the coding sequences of unidentified human genes. XVI. The complete
A/Reference number: A59438
A/Accession: A59438
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-1944 <NAG>
A/Cross-references: UNIPROT:Q9P2C3; GB:BAA92662; PID:g20521912; PIDN:BAA92662.2
R/Ohara, O.; Nagase, T.; Kikuno, R.
submitted to Genbank, January 2000
A/Reference number: B59438
A/Accession: B59438

A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1944 <OHA>

A;Cross-references: GB:BAA92662; PID:g20521912; PIDN:BAA92662.2

Query Match 64.3%; Score 92; DB 2; Length 1944;

Best Local Similarity 64.0%; Pred. No. 6.3e-05;

Matches 16; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MRDLPGHYYETLKLVLGHLKTIADH 25

:|||:|||||:|||||:|

Db 1249 IHDLPPEHHYETLKLFLGAHLKTVAN 1273

RESULT 3

S44876

ZC21.4 protein - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 14-Sep-1994 #sequence_revision 12-May-1995 #text_change 23-Mar-2001

C;Accession: S44876

R;Du, Z.; Waterston, R.

Submitted to the EMBL Data Library, May 1993

A;Description: Sequence of the C. elegans cosmid ZC21.

A;Reference number: S44649

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-733 <DUZ>

A;Cross-references: EMBL:L16695; NID:g289729; PID:g289735

C;Genetics:

A;Introns: 269/3; 551/3; 600/2; 670/3

Query Match

Best Local Similarity 46.9%; Score 67; DB 2; Length 733;

Matches 12; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 1 MRDLPGHYYETLKLVLGHLKTIADH 25

:|||:|||||:|||||:|

Db 198 LRKLPRPHYDTLRLFLVHLSEITKH 222

RESULT 4

E59430

PTPLI-associated RhoGAP protein 1 [imported] - human

C;Species: Homo sapiens (man)

C;Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 09-Jul-2004

C;Accession: E59430

R;Saras, J.; Franzen, P.; Aspenstrom, P.; Hellman, U.; Ginez, L.J.; Heldin, C.-H.

Submitted to Genbank, December 1997

A;Description: Homo sapiens PTPLI-associated RhoGAP 1 (PARG1), mRNA.

A;Reference number: E59430

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1261 <SAR>

A;Cross-references: UNIPROT:O15463; GB:NP_004806; PID:g4758882; PIDN:NP_004806.1

Query Match

Best Local Similarity 45.5%; Score 65; DB 2; Length 1261;

Matches 11; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 MRDLPGHYYETLKLVLGHLKTIADH 25

:|||:|||||:|||||:|

Db 807 LRQLPASNFSLHFLVLHLKRVVDH 831

RESULT 5

T27959

hypothetical protein ZK669.1b - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T27959; T27974

R;Thomas, K.

N;Alternate names: GTPase-activating protein

submitted to the EMBL Data Library, September 1994

A;Reference number: Z20446

A;Accession: T27959

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-634 <WIL>

A;Cross-references: UNIPROT:Q7YSI8; EMBL:Z37093; PIDN:CAA85469.1; GSPDB:GN00020; CESP:ZK669.1a

A;Experimental source: clone ZK669

R;Sims, M.

Submitted to the EMBL Data Library, November 1994

A;Reference number: Z20448

A;Accession: T27974

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-634 <W12>

A;Cross-references: EMBL:Z46812; PIDN:CAA86848.1; GSPDB:GN00020; CESP:ZK669.1b

A;Experimental source: clone ZK675

C;Genetics:

A;Gene: CESP:ZK669.1b

A;Map position: 2

A;Introns: 9/3; 57/2; 158/3; 211/3; 325/2; 474/1; 524/1; 551/3; 563/1

Query Match

Best Local Similarity 44.8%; Score 64; DB 2; Length 634;

Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 2 RDLPGHYETLKLVLGHLKTI A 23

:|||:|||||:|||||:|

Db 303 RKLPHVNYETLKLFLMLHNRVS 324

RESULT 6

T27958

hypothetical protein ZK669.1a - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T27958; T27973

R;Thomas, K.

Submitted to the EMBL Data Library, September 1994

A;Reference number: Z20446

A;Accession: T27958

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-863 <W1L>

A;Cross-references: UNIPROT:Q27267; EMBL:Z37093; PIDN:CAA85468.1; GSPDB:GN00020; CESP:ZK669.1a

A;Experimental source: clone ZK669

R;Sims, M.

Submitted to the EMBL Data Library, November 1994

A;Reference number: Z20448

A;Accession: T27973

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-863 <W12>

A;Cross-references: EMBL:Z46812; PIDN:CAA86847.1; GSPDB:GN00020; CESP:ZK669.1a

A;Experimental source: clone ZK675

C;Genetics:

A;Gene: CESP:ZK669.1a

A;Map position: 2

A;Introns: 62/3; 125/3; 185/3; 211/3; 238/3; 286/2; 387/3; 440/3; 554/2; 703/1; 753/1; 763/1

Query Match

Best Local Similarity 44.8%; Score 64; DB 2; Length 863;

Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 2 RDLPGHYETLKLVLGHLKTI A 23

:|||:|||||:|||||:|

Db 532 RKLPHVNYETLKLFLMLHNRVS 553

RESULT 7

S29128

N-chimerin - rat

N;Alternate names: GTPase-activating protein

C:Species: Rattus norvegicus (Norway rat)
C:Date: 25-Feb-1994 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004
C:Accession: S29128; S25152
R/Lim, H.H.; Michael, G.J.; Smith, P.; Lim, L.; Hall, C.
Biochem. J. 287, 415-422, 1992
A:Title: Developmental regulation and neuronal expression of the mRNA of rat n-chimaerin
A:Reference number: S29128; MUID:93074974; PMID:1445199
A:Accession: S29128
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-334 <LIM>
A:Cross-references: UNIPROT:P30337; EMBL:X67250; NID:G55939; PIDN:CAA47672.1; PID:G55940
R/Lim, H.; Michael, G.J.; Smith, P.; Lim, L.; Hall, C.
Submitted to the EMBL Data Library, July 1992
A:Description: Rat n-chimaerin a p21rac GAP:CDNA sequence developmental regulation and n
A:Reference number: S25152
A:Accession: S25152
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-334 <LIM2>
A:Cross-references: EMBL:X67250; NID:G55939; PIDN:CAA47672.1; PID:G55940
F:81-130/Domain: protein kinase C zinc-binding repeat homology <K22>

Query Match 42.7%; Score 61; DB 2; Length 334;
Best Local Similarity 44.0%; Pred. No. 0.33;
Matches 11; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 MRDLPCHYVETLKFVLVGHKLTIADH 25
Db 258 LRSLLPACHCETLRYLAHLKRVTLH 282
:::|||||:::

RESULT 8
S50658
N:Alternate names: GTPase-activating protein IPL2; protein YER155C
C:Species: Saccharomyces cerevisiae
C:Date: 28-Jan-1995 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
C:Accession: S50658; A55155; A55156
R/Dietrich, F.S.
Submitted to the EMBL Data Library, December 1994
A:Description: The sequence of S. cerevisiae cosmids 8229, 9115, 9132, 9981, and lambda
A:Reference number: S50658
A:Accession: S50658
A:Molecule type: DNA
A:Residues: 1-2167 <DIB>
A:Cross-references: UNIPROT:P39960; EMBL:U18917; NID:G603377; PIDN:AA64682.1; PID:G6033
R/Kim, Y.J.; Francisco, L.; Chen, G.C.; Marcotte, E.; Chan, C.S.M.
J. Cell Biol. 127, 1381-1394, 1994
A:Title: Control of cellular morphogenesis by the Ipl2/Bem2 GTPase-activating protein: P
A:Reference number: A55155; MUID:95050996; PMID:7962097
A:Accession: A55155
A:Molecule type: DNA
A:Residues: 1-2167 <XIM>
A:Cross-references: GB:Z35159; NID:G511136; PIDN:CAA84524.1; PID:G511137
R/Peterson, J.; Zheng, Y.; Bender, L.; Myers, A.; Cerione, R.; Bender, A.
J. Cell Biol. 127, 1395-1406, 1994
A:Title: Interactions between the bud emergence proteins Bem1p and Bem2p and Rho-type G
A:Reference number: A55156; MUID:95050997; PMID:7962098
A:Accession: A55156
A:Molecule type: DNA
A:Residues: 1-2167 <PET>
A:Cross-references: GB:L33832; NID:G499694; PIDN:AAA57132.1; PID:G499695
C:Genetics:
A:Gene: SGD:BEM2; IPL2
A:Cross-references: SGD:S0000957; MIPS:YER155C
A:Map position: 5R
F:2118-2121/Region: GTP-binding NKXD motif

Query Match 40.6%; Score 58; DB 2; Length 2167;
Best Local Similarity 44.0%; Pred. No. 6.5;
Matches 11; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 MRDLPCHYVETLKFVLVGHKLTIADH 25
Db 2090 LQKQTCYYQTLKRIVFHLNKHQH 2114
:::|||||:::

RESULT 9
A59437
K1AA1204 protein [imported] - human
C:Species: Homo sapiens (man)
C:Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 09-Jul-2004
C:Accession: A59437; B59437
R/Nagase, T.; Ishikawa, K.; Kikuno, R.; Hirose, M.; Nomura, N.; Ohara, O.
DNA Res. 6, 337-345, 1999
A:Title: Prediction of the coding sequences of unidentified human genes. XV. The complet
A:Reference number: A59437
A:Accession: A59437
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1445 <NAG>
A:Cross-references: UNIPROT:Q9ULL6; GB:BAA86518; PID:G6382020; PIDN:BAA86518.1
R/Ohara, O.; Nagase, T.; Kikuno, R.
Submitted to GenBank, October 1999
A:Reference number: B59437
A:Accession: B59437
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1445 <OHA>
A:Cross-references: GB:BAA86518; PID:G6382020; PIDN:BAA86518.1

Query Match 39.2%; Score 56; DB 2; Length 1445;
Best Local Similarity 43.5%; Pred. No. 8.3;
Matches 10; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 MRDLPCHYVETLKFVLVGHKLTIADH 23
Db 136 IQELPPSHRYTLELYLRHLAHIA 158
:::|||||:::

RESULT 10
S29956
beta-chimerin - rat
N:Alternate names: GTPase-activating protein
C:Species: Rattus norvegicus (Norway rat)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: A45485; S31398; S29956
R/Leung, T.; How, B.E.; Manser, E.; Lim, L.
J. Biol. Chem. 269, 3813-3816, 1993
A:Title: Germ cell beta-chimaerin, a new GTPase-activating protein for p21rac, is specifi
A:Reference number: A45485; MUID:93179371; PMID:8440677
A:Accession: A45485
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-295 <LEU>
A:Cross-references: UNIPROT:Q03070; EMBL:X69489; NID:G57526; PIDN:CAA49244.1; PID:G57527
A:Experimental source: testis
A:Note: sequence extracted from NCBI backbone (NCBI:P:125731)
R/Leung, T.; How, B.E.; Manser, E.; Lim, L.
Submitted to the EMBL Data Library, November 1992
A:Description: Germ cell beta-chimaerin, a new GTPase-activating protein for p21rac, is
A:Reference number: S31398
A:Accession: S31398
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-295 <LE2>
A:Cross-references: EMBL:X69462; NID:G55816; PID:G55817
F:42-91/Domain: protein kinase C zinc-binding repeat homology <K22>

Query Match 38.5%; Score 55; DB 2; Length 295;
Best Local Similarity 52.6%; Pred. No. 2.2;
Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 4 LPGHYVETLKFVLVGHKLTIADH 22
:::|||||:::

Query Match 38.5%; Score 55; DB 2; Length 295;
Best Local Similarity 52.6%; Pred. No. 2.2;
Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

```
Db 222 LPPAHYETRLYLMIHLKKV 240
RESULT 11
B53764
beta2-chimerin, cerebellar - rat (fragment)
N/Alternate names: GTPase-activating protein
C/Species: Rattus norvegicus (Norway rat)
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 04-Apr-2004
C/Accession: B53764
R/Leung, T.; How, B.E.; Manser, E.; Lim, L.
J. Biol. Chem. 269, 12888-12892, 1994
A/Title: Cerebellar beta2-chimerin, a GTPase-activating protein for p21 Ras-related Rac
A/Reference number: A53764; MUID:94230370; PMID:8175705
A/Accession: B53764
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-443 <LEU>
F:34-114/Domain: SH2 homology <SH2>
F:190-239/Domain: protein kinase C zinc-binding repeat homology <KZ2>
Query Match 38.5%; Score 55; DB 2; Length 443;
Best Local Similarity 52.6%; Pred. No. 3.3;
Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 4 LPGHYETLKLFLVGHLLKTI 22
Db 370 LPPAHYETRLYLMIHLKKV 388
RESULT 12
A53764
beta2-chimerin, cerebellar - human
N/Alternate names: GTPase-activating protein
C/Species: Homo sapiens (man)
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C/Accession: A53764
R/Leung, T.; How, B.E.; Manser, E.; Lim, L.
J. Biol. Chem. 269, 12888-12892, 1994
A/Title: Cerebellar beta2-chimerin, a GTPase-activating protein for p21 Ras-related Rac
A/Reference number: A53764; MUID:94230370; PMID:8175705
A/Accession: A53764
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-466 <LEU>
A/Cross-references: UNIPROT:P52757; GB:L29126; NID:G457229; PIDN:AAA191.1; PID:G457230
F:57-137/Domain: SH2 homology <SH2>
F:213-262/Domain: protein kinase C zinc-binding repeat homology <KZ2>
Query Match 38.5%; Score 55; DB 2; Length 466;
Best Local Similarity 52.6%; Pred. No. 3.5;
Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 4 LPGHYETLKLFLVGHLLKTI 22
Db 393 LPPAHYETRLYLMIHLKKV 411
RESULT 13
T19825
hypothetical protein C38D4.5 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T19825
R/Coles, L.
submitted to the EMBL Data Library, October 1994
A/Reference number: Z19183
A/Accession: T19825
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-837 <WIL>
A/Cross-references: UNIPROT:P46941; EMBL:Z46241; PIDN:CAA86318.1; GSPDB:GN00021; CESP:C3
A/Experimental source: clone C38D4
```

```
C/Genetics:
A/Gene: CESP:C38D4.5
A/Map position: 3
A/Introns: 31/3; 81/2; 137/3; 277/2; 435/3; 507/3; 544/2; 604/1; 646/2; 707/1; 791/3
F:96-133/Domain: WW repeat homology <WWR>
Query Match 38.5%; Score 55; DB 2; Length 837;
Best Local Similarity 50.0%; Pred. No. 6.5;
Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;
QY 4 LPGHYETLKLFLVGHLLKTIADH 25
Db 725 LPNENRETLKWLRLHLNRVASH 746
RESULT 14
T46289
hypothetical protein DKFZp434A1010.1 - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C/Accession: T46289
R/Duesternoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A/Reference number: Z23035
A/Accession: T46289
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-862 <AAA>
A/Cross-references: UNIPROT:Q9NT23; EMBL:AL137579
A/Experimental source: adult testis; clone DKFZp434A1010
C/Genetics:
A/Note: DKFZp434A1010.1
Query Match 38.5%; Score 55; DB 2; Length 862;
Best Local Similarity 40.0%; Pred. No. 6.7;
Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
QY 1 MRDLPGHYETLKLFLVGHLLKTIADH 25
Db 4 IQQLPPPHYRTLYLRLHARMARH 28
RESULT 15
T00705
N-chimerin homolog F25965_3 - human
C/Species: Homo sapiens (man)
C/Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C/Accession: T00705
R/Lamerdin, J.E.; McCready, P.M.; Adamson, A.W.; Burkhardt-Schultz, K.; Garcia, E.; Kyle,
H.; Olsen, A.O.; Carrano, A.V.
submitted to the EMBL Data Library, October 1997
A/Description: Sequence analysis of a 1MB region in 19q13.1.
A/Reference number: Z14199
A/Accession: T00705
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-903 <LAM>
A/Cross-references: UNIPROT:O14560; EMBL:AC002398; NID:G2529398; PIDN:AAB81198.1; PID:G24
C/Genetics:
A/Map position: 19
A/Introns: 17/3; 68/3; 100/2; 148/3; 176/2; 212/2; 261/1; 312/2; 361/1; 513/1
A/Note: F25965_3
Query Match 38.5%; Score 55; DB 2; Length 903;
Best Local Similarity 40.0%; Pred. No. 7.1;
Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
QY 1 MRDLPGHYETLKLFLVGHLLKTIADH 25
Db 167 IQQLPPPHYRTLYLRLHARMARH 191
Search completed: November 10, 2004, 13:40:39
```

Job time : 6.79275 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 12:27:34 ; Search time 31.7478 Seconds
(without alignments)
471.205 Million cell updates/sec

Title: US-10-092-750-31

Perfect score: 143

Sequence: 1 MRDLPGHYYETLKFLVGHKLTIADHR 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|--------------|--------------------|
| 1 | 134 | 93.7 | 735 | 2 Q9P227 | Q9P227 homo sapien |
| 2 | 93 | 65.0 | 1926 | 2 Q6DFG0 | Q6DFG0 xenopus lae |
| 3 | 92 | 64.3 | 1321 | 2 Q8N3A2 | Q8N3A2 homo sapien |
| 4 | 92 | 64.3 | 1902 | 2 Q71M21 | Q71M21 xenopus lae |
| 5 | 92 | 64.3 | 1902 | 2 AAQ04821 | AAQ04821 xenopus l |
| 6 | 92 | 64.3 | 1944 | 2 Q9P2C3 | Q9P2C3 homo sapien |
| 7 | 92 | 64.3 | 1957 | 2 Q8N119 | Q8N119 homo sapien |
| 8 | 88 | 61.5 | 1282 | 2 Q80TD7 | Q80TD7 mus musculu |
| 9 | 88 | 61.5 | 1944 | 2 Q6DFV3 | Q6DFV3 mus musculu |
| 10 | 67 | 46.9 | 1317 | 1 GAP_CABEL | Q91V57 caenorhabdi |
| 11 | 66 | 46.2 | 1266 | 2 Q8CGF1 | Q8CGF1 mus musculu |
| 12 | 65 | 45.5 | 1261 | 2 O15463 | O15463 homo sapien |
| 13 | 64 | 44.8 | 546 | 2 Q7YS18 | Q7YS18 caenorhabdi |
| 14 | 64 | 44.8 | 775 | 2 Q27267 | Q27267 caenorhabdi |
| 15 | 63 | 44.1 | 1337 | 2 Q6PCS4 | Q6PCS4 brachydanio |
| 16 | 63 | 44.1 | 1337 | 2 AAH59184 | AAH59184 brachydan |
| 17 | 62 | 43.4 | 218 | 2 Q81UG4 | Q81UG4 homo sapien |
| 18 | 62 | 43.4 | 322 | 2 Q9BWG3 | Q9BWG3 homo sapien |
| 19 | 62 | 43.4 | 954 | 2 Q86T64 | Q86T64 homo sapien |
| 20 | 62 | 43.4 | 1738 | 2 Q86YL6 | Q86YL6 homo sapien |
| 21 | 62 | 43.4 | 1740 | 2 Q811P8 | Q811P8 mus musculu |
| 22 | 62 | 43.4 | 1770 | 2 Q94820 | Q94820 homo sapien |
| 23 | 61 | 42.7 | 210 | 2 Q9D9B3 | Q9D9B3 mus musculu |
| 24 | 61 | 42.7 | 334 | 1 CHIN MOUSE | Q91V57 mus musculu |
| 25 | 61 | 42.7 | 334 | 1 CHIN RAT | P30337 rattus norv |
| 26 | 61 | 42.7 | 334 | 2 Q8BWU6 | Q8BWU6 mus musculu |
| 27 | 61 | 42.7 | 334 | 2 BAC35853 | BAC35853 mus muscu |
| 28 | 61 | 42.7 | 459 | 2 Q7TQ85 | Q7TQ85 mus musculu |
| 29 | 58 | 40.6 | 332 | 2 Q6GN15 | Q6GN15 xenopus lae |
| 30 | 58 | 40.6 | 826 | 2 Q7PVE8 | Q7PVE8 anopheles g |
| 31 | 58 | 40.6 | 2167 | 1 BEM2_YEAST | P39960 saccharomyc |

RESULT 1

Q9P227 PRELIMINARY; PRT; 735 AA.
AC Q9P227;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE KIAA1501 Protein (Fragment).
GN Name=KIAA1501;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2027482; PubMed=10819331;
RA Nagase T., Kakuno R., Ishikawa K., Hiroseawa M., Ohara O.;
RT Prediction of the coding sequences of unidentified human genes. XVII.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.;
RL DNA Res. 7:143-150(2000).
DR EMBL; AB040934; BAA96025.1; --
DR PIR; A59434; A59434.
DR InterPro; IPR001849; PH.
DR InterPro; IPR01036; PH-related.
DR InterPro; IPR00198; RhoGAP.
DR InterPro; IPR008936; Rho_GAP.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00620; RhoGAP; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00324; RhoGAP; 1.
DR PROSITE; PS50003; PH DOMAIN; 1.
DR PROSITE; PS50238; RHO GAP; 1.
FT NON_TER 1
SQ SEQUENCE 735 AA; 81140 MW; E9981196E1F46D9D CRC64;

Query Match 93.7%; Score 134; DB 2; Length 735;

Best Local Similarity 96.0%; Pred.No. 1.2e-10;

Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRDLPGHYYETLKFLVGHKLTIADH 25

Db 612 IRDLPGHYYETLKFLVGHKLTIADH 636

RESULT 2

Q6DFG0 PRELIMINARY; PRT; 1926 AA.
AC Q6DFG0;
DT 01-OCT-2004 (Tremblrel. 28, Created)
DT 01-OCT-2004 (Tremblrel. 28, Last sequence update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Oocytes;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulláhy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [2]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Oocytes;
 RX MEDLINE=22341132; PubMed=12454917;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative";
 RL Dev. Dyn. 225:384-391(2002).
 [3]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Oocytes;
 RX Klein S., Strausberg R.;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC076778; AAH76778.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 1926 AA; 214711 MW; FA7FCF57681DD32 CRC64;
 Query Match 65.0%; Score 93; DB 2; Length 1926;
 Best Local Similarity 69.6%; Pred. No. 0.00028;
 Matches 16; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 QY 3 DLPGHYETLKLVLGHLKTIADH 25
 |||||:|||||:|||||:
 DB 1244 DLPDHYETLKLVLGHLKTIADN 1266
 |||||:|||||:|||||:
 RESULT 3
 Q8N3A2 PRELIMINARY; PRT; 1321 AA.
 ID Q8N3A2
 AC Q8N3A2
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein DKFZP613J031 (Fragment).
 GN Name=DKFZP613J031;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Amalgam;
 RX Ottenwaelder B., Obermaier B., Mewes H.W., Weil B., Wiemann S.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL834495; CAD39153.1; -;
 DR InterPro; IPR001849; PH related.
 DR SMART; SM001036; PH; 1.
 DR InterPro; IPR000198; RhoGAP.
 DR InterPro; IPR008936; Rho GAP.
 DR InterPro; IPR001605; SpeGtrn_PH.
 DR Pfam; PF00169; PH; 1.
 DR Pfam; PF00620; RhoGAP; 1.
 DR PRINTS; PRO0683; SPECTRINPH.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00324; RhoGAP; 1.
 DR PROSITE; PS50003; PH DOMAIN; 1.
 DR PROSITE; PS50238; RHO GAP; 1.
 KW Hypothetical protein.
 FT NOETHER 1
 SQ SEQUENCE 1321 AA; 146862 MW; A190E9B171EEF015 CRC64;
 Query Match 64.3%; Score 92; DB 2; Length 1321;
 Best Local Similarity 64.0%; Pred. No. 0.00027;
 Matches 16; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 QY 1 MDLPGHYETLKLVLGHLKTIADH 25
 :|||||:|||||:|||||:
 DB 626 IHDUPHETLKLVLGHLKTIADN 650
 :|||||:|||||:|||||:
 RESULT 4
 Q71M21 PRELIMINARY; PRT; 1902 AA.
 ID Q71M21
 AC Q71M21;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Rho-GTPase activating protein.
 GN Name=rgap;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=22697565; PubMed=12711552;
 RX Kim J., Shim S., Choi S.C., Han J.K.;
 RT "A putative Xenopus Rho-GTPase activating protein (XrGAP) gene is
 RT expressed in the notochord and brain during the early embryogenesis.";
 RL Gene Expr. Patterns 3:219-223(2003).
 CC -1- SIMILARITY: Contains 1 PDZ/DHR domain.
 DR EMBL; AF462392; AAQ04821.1; -;
 DR InterPro; IPR001478; PDZ.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR01036; PH related.
 DR InterPro; IPR000198; RhoGAP.
 DR InterPro; IPR008936; Rho GAP.
 DR Pfam; PF00595; PDZ; 1.
 DR Pfam; PF00169; PH; 1.
 DR Pfam; PF00620; RhoGAP; 1.
 DR SMART; SM00228; PDZ; 1.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00324; RhoGAP; 1.
 DR PROSITE; PS50106; PDZ; 1.
 DR PROSITE; PS50003; PH DOMAIN; 1.
 DR PROSITE; PS50238; RHO GAP; 1.
 SQ SEQUENCE 1902 AA; 211461 MW; C61B9567A556B01E CRC64;
 Query Match 64.3%; Score 92; DB 2; Length 1902;
 Best Local Similarity 72.7%; Pred. No. 0.00039;
 Matches 16; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 3 DLPGHYETLKLVLGHLKTIAD 24
 |||||:|||||:|||||:
 DB 1221 DLPDHYETLKLVLGHLKTIAD 1242
 |||||:|||||:|||||:


```

RESULT 5
AAQ04821 PRELIMINARY; PRT; 1902 AA.
AC AAQ04821;
DT 20-MAY-2004 (TrEMBLrel. 27, Created)
DT 20-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DE Rho-GTPase activating protein.
GN RGAAP.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22597565; PubMed=12711552;
RA Kim J., Shim S., Choi S.C., Han J.K.;
RT "A putative Xenopus Rho-GTPase activating protein (XrGAP) gene is
RT expressed in the notochord and brain during the early embryogenesis.";
RL Gene Expr. Patterns 3:219-223(2003).
DR EMBL; AF462392; AAQ04821.1; -.
SQ SEQUENCE 1902 AA; 211461 MW; C61B9567A556E01E CRC64;

Query Match 64.3%; Score 92; DB 2; Length 1902;
Best Local Similarity 72.7%; Pred.No. 0.00039; Mismatches 3; Indels 0; Gaps 0;
Matches 16; Conservative 3;

QY 3 DLPGHYETLKFLVGHKLTIAD 24
Db 1221 DLPDHYETLKFLVGHKLTIAD 1242

RESULT 6
Q9P2C3 PRELIMINARY; PRT; 1944 AA.
AC Q9P2C3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE KIAA1424 protein (Fragment).
GN Name=KIAA1424;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
RA Nagase T., Kikuno R., Ishikawa K., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 7:65-73(2000).
CC -1- SIMILARITY: Contains 1 PDZ/DHR domain.
DR EMBL; AB037845; BAA92662.2; -.
DR PIR; A59438; A59438.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR001478; PDZ.
DR InterPro; IPR001849; PH.
DR InterPro; IPR011036; PH-related.
DR InterPro; IPR000198; RhoGAP.
DR InterPro; IPR008936; Rho_GAP.
DR Pfam; PF00595; PDZ; 1.
DR Pfam; PF00169; PH; 1.
DR SMART; SM00228; PDZ; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00324; RhoGAP; 1.
DR PROSITE; PS50106; PDZ; 1.
DR PROSITE; PS50003; PH DOMAIN; 1.
DR PROSITE; PS50238; RHO_GAP; 1.
SQ SEQUENCE 1957 AA; 217325 MW; 57C90CE68B0FA164 CRC64;

Query Match 64.3%; Score 92; DB 2; Length 1957;
Best Local Similarity 64.0%; Pred.No. 0.0004;
Matches 16; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 MRDLPGHYYETLKFLVGHKLTIADH 25
Db 1262 IHDLPENHYETLKFLSAHLKTVAN 1296

RESULT 8
Q80TD7 PRELIMINARY; PRT; 1262 AA.
AC Q80TD7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE KIAA1424 protein (Fragment).
GN Name=Arhgap21; Synonyms=mKIAA1424;

```

Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smalley D.S., Schnerch A., Schein J.E., Jones S.J., Marra M.A., "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.", Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

[2]

RN
RP
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Head;
R Strausberg R.;
RRL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC076629; AAH76629.1 -;
RSQ SEQUENCE 1944 AA; 215741 MW; A7F832ECD3DBC081 CRC64;

Query Match 61.5%; Score 88; DB 2; Length 1944;
Best_Local_Similarity 60.08; Pred. No. 0.0015;
Matches 15; Conservative 6; Mismatches 4; Indels 0; Gaps

QY 1 MRDLPGHYETLKFLVGLHLTIADH 25
: |||:||||| |||||:
Db 1256 IHDLPFHFFETKFLSAHLKTVAEN 1280

RESULT 10
GAP_CAEEL STANDARD; PRT; 1317 AA.
ID GAP_CAEEL
AC P34288; P34589; Q8ST07; Q9TY64;
DT 01-FEB-1994 (Rel. 28, Created)
DI 10-OCT-2003 (Rel. 42, Last sequence update)
DS 05-JUL-2004 (Rel. 44, Last annotation update)
DE GPase-activating protein GAP (CeGAP).
GN Name=gap; ORSNames=C04D8.1, ZC21.4/ZC21.5;
OS Caenorhabditis elegans
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OX Rhabditiidae; Pelicerinae; Caenorhabditis.
NCBI_TaxID=6239;
LN (1)
RP SEQUENCE FROM N.A., AND FUNCTION.
RC STRAIN=Bristol N2;
RX MEDLINE=94117490; PubMed=8288633;
RA Chen W., Blanc J., Lim L.;
RT "Characterization of a promiscuous GTPase-activating protein that has
RL a bor-related domain from *Caenorhabditis elegans*.";
RN J. Biol. Chem. 269:820-823(1994).
RS [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Bonfield J., Burton J., Connell M., Baynes C., Berks M., Coulson A.,
RA Craxton L., Dear S., Du Z., Durbin R., Favella A., Fraser A.,
RA Fulton M., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkhen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonnenhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Watkinson R., Watson A., Weinstein L., Wilkinson-Spratt J.,
RA Wohlsman P.;
RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of *C.*
RT elegans.";
RL Nature 368:32-38(1994).
RS [3]
RP REVISIONS.
RC Watson R.;
RRL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: GPase-activating protein for members of the Rho
CC subfamily including Rac1, RhoA and cdc42 and other Ras-related
CC subfamilies including let-60.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Contains 1 PH domain.
CC -!- SIMILARITY: Contains 1 Rho-GAP domain.

```

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; U02289; AAA18934.1; ALT_INIT.
EMBL; L16687; AAK71357.2; -.
PIR; S44746; S44746.
PIR; S44877; S44877.
HSSP; Q07960; IRGP.
WormPep; CO4D8.1; CE33028.
GO; GO:0005737; C:cytoplasm; NAS.
GO; GO:0005099; F:Ras GTPase activator activity; IDA.
GO; GO:0005100; F:Rho GTPase activator activity; IDA.
GO; GO:0007265; P:Ras protein signal transduction; IDA.
GO; GO:0007266; P:Rho protein signal transduction; IDA.
InterPro; IPR001849; PH.
InterPro; IPR011036; PH related.
InterPro; IPR000198; RhoGAP.
InterPro; IPR008936; RhoGAP.
Pfam; PF00169; PH; 1.
Pfam; PF00620; RhoGAP; 1.
SMART; SM00233; PH; 1.
SMART; SM00324; RhoGAP; 1.
PROSITE; PS50003; PH DOMAIN; 1.
PROSITE; PS50238; RHO GAP; 1.
KW GTPase activation. 264
FT DOMAIN 106 Ser-rich.
FT DOMAIN 311 PH.
FT DOMAIN 496 513 Glu-rich.
FT DOMAIN 660 858 Rho-GAP.
FT DOMAIN 1064 1088 Ala-rich.
FT CONFLICT 276 276 L -> I (in Ref. 1).
SQ SEQUENCE 1317 AA; 142069 MW; 5871763017001A39 CRC64;

Query Match 46.9%; Score 67; DB 1; Length 1317;
Best Local Similarity 48.0%; Pred. No. 1.2;
Matches 12; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 MRDLPGHYYETLKLFLVGHKLTIADH 25
Db 782 LRKLPRPHYDTLRLVHLSLSEITKH 806

RESULT 11
ID QBCGF1 PRELIMINARY; PRT; 1266 AA.
AC QBCGF1
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PTP1L-associated RhoGAP 1.
GN Name=El3001701Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N;
RC TISSUE=Mammary tumor. Metallothionien-TGF alpha model. 10 month old
RC virgin mouse. Taken by biopsy;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

```

```

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny K.C., Sodergren E.U., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N;
RC TISSUE=Mammary tumor. Metallothionien-TGF alpha model. 10 month old
RC virgin mouse. Taken by biopsy;
RA Strausberg R.;
RL EMBL; BC040387; AAH40387.1; -.
DR HSSP; Q98935; 1F7C.
DR MGD; MGI:2443818; BL30017101Rik.
DR CO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR000198; RhoGAP.
DR InterPro; IPR008936; RhoGAP.
DR Pfam; PF00130; Cl 1; 1.
DR Pfam; PF00620; RhoGAP; 1.
DR SMART; SM00109; Cl; 1.
DR SMART; SM00324; RhoGAP; 1.
DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; UNKNOWN_1.
DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE; PS50238; RHO GAP; 1.
DR PROSITE; PS50238; RHO GAP; 1.
SQ SEQUENCE 1266 AA; 142339 MW; 96687B52C3D229D2 CRC64;

Query Match 46.2%; Score 66; DB 2; Length 1266;
Best Local Similarity 36.0%; Pred. No. 1.6;
Matches 9; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 1 MRDLPGHYYETLKLFLVGHKLTIADH 25
Db 809 LRQLPASHFNSLHLYLHRLRWVDH 833

RESULT 12
ID O15463 PRELIMINARY; PRT; 1261 AA.
AC O15463
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PTP1L-associated RhoGAP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=97450957; PubMed=9305890;
RA Saras J., Franzen P., Aspenstrom P., Hellman U., Genez L.J.,
RA Heldin C.H.;
RT "A novel GTPase-activating protein for Rho interacts with a PDZ domain
of the protein-tyrosine phosphatase PTP1L.";
J. Biol. Chem. 272:24333-24338(1997).
DR EMBL; U09920; AAB81012.1; -.
DR PIR; E59430; E59430.
DR HSSP; Q98935; 1F7C.
DR GO; GO:0005100; F:Rho GTPase activator activity; TAS.
DR GO; GO:0007266; P:Rho protein signal transduction; TAS.
DR InterPro; IPR002219; DAG_PE-bind.

```

Page 6

[illegible]

Search completed: November 10, 2004, 13:38:30
Job time : 33.7478 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 12:32:37 ; Search time 14.5078 Seconds
(without alignments)
191.991 Million cell updates/sec

Title: US-10-092-750-32
Perfect score: 232
Sequence: 1 CGRMEDIPCSRVGHIYKY.....YKVPAGVSLARNLKRVDWM 42

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/iaa/PCUTUS_COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 217.5 | 93.8 | 506 | 4 | US-09-795-926-31 |
| 2 | 217.5 | 93.8 | 535 | 4 | US-09-795-926-29 |
| 3 | 217.5 | 93.8 | 603 | 4 | US-09-795-926-43 |
| 4 | 217.5 | 93.8 | 631 | 4 | US-09-795-926-41 |
| 5 | 120.5 | 51.9 | 149 | 4 | US-09-973-457-6 |
| 6 | 117.5 | 50.6 | 230 | 4 | US-09-270-767-61773 |
| 7 | 117.5 | 50.6 | 313 | 4 | US-09-270-767-46209 |
| 8 | 116.5 | 50.2 | 552 | 4 | US-10-140-002-196 |
| 9 | 110.5 | 47.6 | 101 | 4 | US-09-270-767-33441 |
| 10 | 110.5 | 47.6 | 101 | 4 | US-09-270-767-48658 |
| 11 | 109.5 | 47.2 | 517 | 2 | US-08-967-508-19 |
| 12 | 109.5 | 47.2 | 517 | 3 | PCT-US94-02552-19 |
| 13 | 109.5 | 47.2 | 517 | 5 | US-08-967-506-19 |
| 14 | 109.5 | 47.2 | 559 | 2 | US-08-967-508-9 |
| 15 | 109.5 | 47.2 | 559 | 3 | US-08-967-506-9 |
| 16 | 109.5 | 47.2 | 559 | 5 | PCT-US94-02552-9 |
| 17 | 109 | 47.2 | 633 | 2 | US-08-648-298-2 |
| 18 | 106.5 | 45.9 | 73 | 4 | US-09-270-767-33350 |
| 19 | 106.5 | 45.9 | 73 | 4 | US-09-270-767-48567 |
| 20 | 96 | 41.4 | 560 | 4 | US-09-217-306B-10 |
| 21 | 96 | 41.4 | 578 | 4 | US-09-217-306B-2 |
| 22 | 95.5 | 41.2 | 603 | 4 | US-09-973-457-2 |
| 23 | 88.5 | 38.1 | 638 | 3 | US-09-347-488-2 |
| 24 | 88.5 | 38.1 | 639 | 3 | US-09-376-856-2 |
| 25 | 87 | 37.5 | 302 | 4 | US-09-270-767-33733 |
| 26 | 57 | 24.6 | 289 | 4 | US-09-270-767-45334 |
| 27 | 53.5 | 23.1 | 386 | 2 | US-08-858-052-4 |

| | | | | | | |
|----|------|------|------|---|----------------------|-------------------|
| 28 | 53.5 | 23.1 | 386 | 3 | US-09-200-284-4 | Sequence 4, Appli |
| 29 | 53 | 22.8 | 576 | 4 | US-09-252-991A-26023 | Sequence 26023, A |
| 30 | 52 | 22.4 | 140 | 4 | US-09-710-279-2610 | Sequence 2610, Ap |
| 31 | 52 | 22.4 | 313 | 4 | US-09-538-992-722 | Sequence 722, App |
| 32 | 52 | 22.4 | 389 | 3 | US-09-307-621-2 | Sequence 2, Appli |
| 33 | 52 | 22.4 | 602 | 3 | US-09-710-279-2362 | Sequence 2362, Ap |
| 34 | 52 | 22.4 | 612 | 3 | US-09-134-001C-4028 | Sequence 4028, Ap |
| 35 | 50.5 | 21.8 | 806 | 4 | US-09-328-352-7160 | Sequence 7160, Ap |
| 36 | 50.5 | 21.8 | 1016 | 4 | US-09-252-991A-18914 | Sequence 18914, A |
| 37 | 50 | 21.6 | 82 | 4 | US-09-270-767-36846 | Sequence 36846, A |
| 38 | 50 | 21.6 | 82 | 4 | US-09-270-767-52063 | Sequence 52063, A |
| 39 | 50 | 21.6 | 352 | 2 | US-08-472-534-6 | Sequence 6, Appli |
| 40 | 50 | 21.6 | 378 | 4 | US-09-583-110-4647 | Sequence 4647, Ap |
| 41 | 50 | 21.6 | 390 | 4 | US-09-634-238-331 | Sequence 331, Ap |
| 42 | 50 | 21.6 | 496 | 2 | US-08-511-485-10 | Sequence 10, Appl |
| 43 | 50 | 21.6 | 496 | 3 | US-09-212-971-10 | Sequence 10, Appl |
| 44 | 50 | 21.6 | 496 | 3 | US-08-800-929A-10 | Sequence 10, Appl |
| 45 | 50 | 21.6 | 496 | 3 | US-09-617-053A-10 | Sequence 10, Appl |

ALIGNMENTS

RESULT 1
US-09-795-926-31
; Sequence 31, Application US/09795926
; Patent No. 6555669
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieker, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 506
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-795-926-31

Query Match 93.8%; Score 217.5; DB 4; Length 506;
Best Local Similarity 95.3%; Pred. No. 2.4e-24;
Matches 41; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 CGRMEDIPCSRVGHIYKYVPYKVPAGVSLARNLKRVDWM 42
DB 259 CGRMEDIPCSRVGHIYKYVPYKVPAGVSLARNLKRVDWM 301

RESULT 2
US-09-795-926-29
; Sequence 29, Application US/09795926
; Patent No. 6555669
; GENERAL INFORMATION:

```

; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/09/795,926
; CURRENT FILING DATE: 2001-02-28
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 535
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-795-926-29

Query Match      93.8%; Score 217.5; DB 4; Length 535;
Best Local Similarity 95.3%; Pred. No. 2.9e-24;
Matches 41; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 1 CGGRMEDIPCSRVGHIYKVPYKVPAGVSLARNLKRKRVAD-WM 42
      |||
Db 288 CGGRMEDIPCSRVGHIYKVPYKVPAGVSLARNLKRKRVAEVWM 330
      |||

RESULT 3
US-09-795-926-43
; Sequence 43, Application US/09795926
; Patent No. 6555669
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/09/795,926
; CURRENT FILING DATE: 2001-02-28
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 603
; TYPE: PRT

```

```

; ORGANISM: homo sapiens
US-09-795-926-43

Query Match      93.8%; Score 217.5; DB 4; Length 603;
Best Local Similarity 95.3%; Pred. No. 2.9e-24;
Matches 41; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 1 CGGRMEDIPCSRVGHIYKVPYKVPAGVSLARNLKRKRVAD-WM 42
      |||
Db 356 CGGRMEDIPCSRVGHIYKVPYKVPAGVSLARNLKRKRVAEVWM 398
      |||

RESULT 4
US-09-795-926-41
; Sequence 41, Application US/09795926
; Patent No. 6555669
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/09/795,926
; CURRENT FILING DATE: 2001-02-28
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 631
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-795-926-41

Query Match      93.8%; Score 217.5; DB 4; Length 631;
Best Local Similarity 95.3%; Pred. No. 3.1e-24;
Matches 41; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 1 CGGRMEDIPCSRVGHIYKVPYKVPAGVSLARNLKRKRVAD-WM 42
      |||
Db 385 CGGRMEDIPCSRVGHIYKVPYKVPAGVSLARNLKRKRVAEVWM 427
      |||

RESULT 5
US-09-973-457-6
; Sequence 6, Application US/09973457
; Patent No. 6703230
; GENERAL INFORMATION:
; APPLICANT: Kapseller-Libermann, Rosana
; TITLE OF INVENTION: 47174, A NOVEL HUMAN GLYCOSYLTRANSFERASE
; FILE REFERENCE: 10448-099001
; CURRENT APPLICATION NUMBER: US/09/973,457
; CURRENT FILING DATE: 2001-10-09
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6

```



```

; LENGTH: 149
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus sequence
US-09-973-457-6

Query Match          51.9%; Score 120.5; DB 4; Length 149;
Best Local Similarity 51.1%; Pred. No. 2.5e-10;
Matches 24; Conservative 7; Mismatches 11; Indels 5; Gaps 2;

Qy 1 CGGRMEDIPCSRGVGHYRKVPYKVPAGVS----LARNLKRVD-WM 42
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 52 CGGRLEIVPCSHVGHVFRKSPYTFPGKSGKGVISRTVRAEVM 98

RESULT 6
US-09-270-767-61773
; Sequence 61773, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61773
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-140-002-196

Query Match          50.6%; Score 117.5; DB 4; Length 290;
Best Local Similarity 53.3%; Pred. No. 1.6e-09;
Matches 24; Conservative 4; Mismatches 14; Indels 3; Gaps 2;

Qy 1 CGGRMEDIPCSRGVGHYRKVPYKVPAGVS--LARNLKRVD-WM 42
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 12 CGGRVEISPCSHVGHVFRSSTPTTFPGMGSEVLTDNLARAATVM 56

RESULT 7
US-09-270-767-46209
; Sequence 46209, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46209
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-140-002-196

Query Match          50.6%; Score 117.5; DB 4; Length 313;
Best Local Similarity 53.3%; Pred. No. 1.7e-09;
Matches 24; Conservative 4; Mismatches 14; Indels 3; Gaps 2;

Qy 1 CGGRMEDIPCSRGVGHYRKVPYKVPAGVS--LARNLKRVD-WM 42
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 12 CGGRVEISPCSHVGHVFRSSTPTTFPGMGSEVLTDNLARAATVM 56

RESULT 8
US-10-140-002-196
; Sequence 196, Application US/10140002

```

```

; Patent No. 6725730
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Wary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C59
; CURRENT APPLICATION NUMBER: US/10/140,002
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 196
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-002-196

Query Match          50.2%; Score 116.5; DB 4; Length 552;
Best Local Similarity 51.1%; Pred. No. 4.7e-09;
Matches 23; Conservative 7; Mismatches 12; Indels 3; Gaps 2;

Qy 1 CGGRMEDIPCSRGVGHYRKVPYKVPAG--VSLARNLKRVD-WM 42
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 319 CGGLEIIVPCSRGVGHVFRKHPYFPDGNANTYINKTKRTAEVWM 363

RESULT 9
US-09-270-767-33441
; Sequence 33441, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33441
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-270-767-33441

Query Match          47.6%; Score 110.5; DB 4; Length 101;
Best Local Similarity 48.9%; Pred. No. 5.1e-09;
Matches 22; Conservative 7; Mismatches 13; Indels 3; Gaps 2;

Qy 1 CGGRMEDIPCSRGVGHYRKVPYKVPAGVS--LARNLKRVD-WM 42
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 33 CGGILEIIPCSRGVGHVFRDKSPYTFPGGVAKIVLHNAARVAEVL 77

RESULT 10
US-09-270-767-48658
; Sequence 48658, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:

```



```

; STREET: 301 Henrietta Street
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: USA
; ZIP: 49001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PC7/US94/02552
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Darnley Jr., James D.
; REGISTRATION NUMBER: 33,673
; REFERENCE/DOCKET NUMBER: 4755.P CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 616-833-8897
; TELEFAX: 616-833-8897
; TELEX: 224401
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 517 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US94-02552-19
;
; Query Match 47.2%; Score 109.5; DB 5; Length 517;
; Best Local Similarity 44.4%; Pred. No. 5e-08;
; Matches 20; Conservative 13; Indels 3; Gaps 2;
;
; QY 1 CGGEMDIPCSRGVGHVYKVPKPVAGVS--LARNLKRVD-WM 42
; DB 288 CGGTLEIVTCSHVGHVFRKATPTFFPGTGQIINKNNRLAEVWM 332
;
; RESULT 14
; US-08-967-508-9
; Sequence 9, Application US/08967508
; Patent No. 5910570
; GENERAL INFORMATION:
; APPLICANT: The Upjohn Company
; APPLICANT: FOR U.S. PURPOSES ONLY: Elhammer, Ake P. and Homa, Fred L.
; TITLE OF INVENTION: A Cloned DNA Encoding a UDP-GalNac:
; TITLE OF INVENTION: Polypeptide, N-Acetylglactosaminyltransferase
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pharmacia & Upjohn Company, Intellectual
; ADDRESSEE: Property Legal Services
; STREET: 301 Henrietta Street
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: USA
; ZIP: 49001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,508
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/602,830
; FILING DATE: 13 No. 5910570ember 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Darnley Jr., James D.
; REGISTRATION NUMBER: 33,673
; REFERENCE/DOCKET NUMBER: 4755.P CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 616-833-8897
; TELEFAX: 616-833-8897
; TELEX: 224401
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 559 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-967-506-9
;
; Query Match 47.2%; Score 109.5; DB 3; Length 559;
; Best Local Similarity 44.4%; Pred. No. 5.4e-08;
; Matches 20; Conservative 9; Mismatches 13; Indels 3; Gaps 2;
;
; QY 1 CGGEMDIPCSRGVGHVYKVPKPVAGVS--LARNLKRVD-WM 42
; DB 330 CGGTLEIVTCSHVGHVFRKATPTFFPGTGQIINKNNRLAEVWM 374
;
; RESULT 15
; US-08-967-506-9
; Sequence 9, Application US/08967506
; Patent No. 6096512
; GENERAL INFORMATION:
; APPLICANT: The Upjohn Company
; APPLICANT: FOR U.S. PURPOSES ONLY: Elhammer, Ake P. and Homa, Fred L.
; TITLE OF INVENTION: A Cloned DNA Encoding a UDP-GalNac:
; TITLE OF INVENTION: Polypeptide, N-Acetylglactosaminyltransferase
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pharmacia & Upjohn Company, Intellectual
; ADDRESSEE: Property Legal Services
; STREET: 301 Henrietta Street
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: USA
; ZIP: 49001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,506
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/602,830
; FILING DATE: 13 No. 6096512ember 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Darnley Jr., James D.
; REGISTRATION NUMBER: 33,673
; REFERENCE/DOCKET NUMBER: 4755.P CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 616-833-8897
; TELEFAX: 616-833-8897
; TELEX: 224401
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 559 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-967-506-9
;
; Query Match 47.2%; Score 109.5; DB 3; Length 559;
; Best Local Similarity 44.4%; Pred. No. 5.4e-08;
; Matches 20; Conservative 9; Mismatches 13; Indels 3; Gaps 2;

```

Oy 1 CGGRMEDIPCSRYGHIYRYKVPYKVPAGVS--LARNLKRVD-WM 42
Db 330 CGGTLEIVTCSHVGVFRKATPVTFPGGTGQIINKNRRRLAEVWM 374

Search completed: November 10, 2004, 13:44:05
Job time : 15.5078 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 16:36:12 ; Search time 45.1192 Seconds
(without alignments)
328.807 Million cell updates/sec

Title: US-10-092-750-32

Perfect score: 232

Sequence: 1 CGGRMEDIPCSRGVGHYRK.....YKVPAGVSLARNLKRVDWM 42

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|------------------|
| 1 | 232 | 100.0 | 42 | 14 | US-10-092-750-32 |
| 2 | 217.5 | 93.8 | 506 | 9 | US-09-795-926-31 |
| 3 | 217.5 | 93.8 | 506 | 14 | US-10-364-774-31 |
| 4 | 217.5 | 93.8 | 535 | 9 | US-09-795-926-29 |
| 5 | 217.5 | 93.8 | 535 | 14 | US-10-364-774-29 |
| 6 | 217.5 | 93.8 | 603 | 9 | US-09-795-926-43 |
| 7 | 217.5 | 93.8 | 603 | 13 | US-10-001-851-2 |
| 8 | 217.5 | 93.8 | 603 | 14 | US-10-364-774-43 |
| 9 | 217.5 | 93.8 | 631 | 9 | US-09-795-926-41 |
| 10 | 217.5 | 93.8 | 631 | 14 | US-10-364-774-41 |
| 11 | 200.5 | 86.4 | 276 | 14 | US-10-292-896-3 |
| 12 | 137 | 59.1 | 62 | 9 | US-09-924-358-35 |
| 13 | 137 | 59.1 | 62 | 14 | US-10-133-709-9 |

| | | | | | | |
|----|-------|------|-----|----|--------------------|-------------------|
| 14 | 137 | 59.1 | 62 | 15 | US-10-410-764-35 | Sequence 35, Appl |
| 15 | 128.5 | 55.4 | 539 | 13 | US-10-087-192-1989 | Sequence 1989, Ap |
| 16 | 128.5 | 55.4 | 558 | 14 | US-10-292-896-66 | Sequence 66, Appl |
| 17 | 128.5 | 55.4 | 558 | 15 | US-10-433-256-6 | Sequence 6, Appl |
| 18 | 128.5 | 55.4 | 571 | 13 | US-10-087-192-1992 | Sequence 1992, Ap |
| 19 | 120.5 | 51.9 | 149 | 9 | US-09-973-457-6 | Sequence 6, Appl |
| 20 | 120.5 | 51.9 | 149 | 14 | US-10-184-648-47 | Sequence 47, Appl |
| 21 | 120.5 | 51.9 | 149 | 15 | US-10-713-345-6 | Sequence 6, Appl |
| 22 | 120.5 | 51.9 | 166 | 13 | US-10-074-527-7 | Sequence 7, Appl |
| 23 | 116.5 | 50.2 | 552 | 9 | US-09-815-028-2 | Sequence 2, Appl |
| 24 | 116.5 | 50.2 | 552 | 14 | US-10-028-072-196 | Sequence 196, App |
| 25 | 116.5 | 50.2 | 552 | 14 | US-10-140-808-196 | Sequence 196, App |
| 26 | 116.5 | 50.2 | 552 | 14 | US-10-121-049-196 | Sequence 196, App |
| 27 | 116.5 | 50.2 | 552 | 14 | US-10-123-904-196 | Sequence 196, App |
| 28 | 116.5 | 50.2 | 552 | 14 | US-10-140-470-196 | Sequence 196, App |
| 29 | 116.5 | 50.2 | 552 | 14 | US-10-175-746-196 | Sequence 196, App |
| 30 | 116.5 | 50.2 | 552 | 14 | US-10-176-918-196 | Sequence 196, App |
| 31 | 116.5 | 50.2 | 552 | 14 | US-10-176-921-196 | Sequence 196, App |
| 32 | 116.5 | 50.2 | 552 | 14 | US-10-137-865-196 | Sequence 196, App |
| 33 | 116.5 | 50.2 | 552 | 14 | US-10-140-474-196 | Sequence 196, App |
| 34 | 116.5 | 50.2 | 552 | 14 | US-10-142-431-196 | Sequence 196, App |
| 35 | 116.5 | 50.2 | 552 | 14 | US-10-143-114-196 | Sequence 196, App |
| 36 | 116.5 | 50.2 | 552 | 14 | US-10-140-002-196 | Sequence 196, App |
| 37 | 116.5 | 50.2 | 552 | 14 | US-10-142-419-196 | Sequence 196, App |
| 38 | 116.5 | 50.2 | 552 | 14 | US-10-123-262-196 | Sequence 196, App |
| 39 | 116.5 | 50.2 | 552 | 14 | US-10-142-423-196 | Sequence 196, App |
| 40 | 116.5 | 50.2 | 552 | 14 | US-10-121-050-196 | Sequence 196, App |
| 41 | 116.5 | 50.2 | 552 | 14 | US-10-141-755-196 | Sequence 196, App |
| 42 | 116.5 | 50.2 | 552 | 14 | US-10-143-032-196 | Sequence 196, App |
| 43 | 116.5 | 50.2 | 552 | 14 | US-10-123-108-196 | Sequence 196, App |
| 44 | 116.5 | 50.2 | 552 | 14 | US-10-123-236-196 | Sequence 196, App |
| 45 | 116.5 | 50.2 | 552 | 14 | US-10-123-261-196 | Sequence 196, App |

ALIGNMENTS

RESULT 1
US-10-092-750-32
; Sequence 32, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-XL
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-32

Query Match 100.0%; Score 232; DB 14; Length 42;
Best Local Similarity 100.0%; Pred. No. 5e-25;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGRMEDIPCSRGVGHYRKVPYKVPAGVSLARNLKRVDWM 42
Db 1 CGGRMEDIPCSRGVGHYRKVPYKVPAGVSLARNLKRVDWM 42

RESULT 2
US-09-795-926-31
; Sequence 31, Application US/09795926
; Patent No. US2002009486A1
; GENERAL INFORMATION:

```

; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/09/795,926
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 506
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-795-926-31

Query Match          93.8%; Score 217.5; DB 9; Length 506;
Best Local Similarity 95.3%; Pred. No. 8.2e-22;
Matches 41; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy      1  CGGRMEDIPCSRGVGHYRKVPYKVPAGVSLARNLKRKVVAD-WM 42
Db      259 CGGRMEDIPCSRGVGHYRKVPYKVPAGVSLARNLKRKVVAD-WM 301

RESULT 3
US-10-364-774-31
; Sequence 31, Application US/10364774
; Publication No. US20030144497A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/10/364,774
; CURRENT FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31

```

```

; LENGTH: 506
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-364-774-31

Query Match          93.8%; Score 217.5; DB 14; Length 506;
Best Local Similarity 95.3%; Pred. No. 8.2e-22;
Matches 41; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy      1  CGGRMEDIPCSRGVGHYRKVPYKVPAGVSLARNLKRKVVAD-WM 42
Db      259 CGGRMEDIPCSRGVGHYRKVPYKVPAGVSLARNLKRKVVAD-WM 301

RESULT 4
US-09-795-926-29
; Sequence 29, Application US/09795926
; Patent No. US20020098486A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/09/795,926
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 535
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-795-926-29

Query Match          93.8%; Score 217.5; DB 9; Length 535;
Best Local Similarity 95.3%; Pred. No. 8.8e-22;
Matches 41; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy      1  CGGRMEDIPCSRGVGHYRKVPYKVPAGVSLARNLKRKVVAD-WM 42
Db      288 CGGRMEDIPCSRGVGHYRKVPYKVPAGVSLARNLKRKVVAD-WM 330

RESULT 5
US-10-364-774-29
; Sequence 29, Application US/10364774
; Publication No. US20030144497A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.

```

APPLICANT: Hu, Yi
APPLICANT: Kieke, James Alvin
APPLICANT: Potter, David George
TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
FILE REFERENCE: LEX-0144-USA
CURRENT FILING DATE: 2003-02-11
PRIOR FILING DATE: 2001-02-28
PRIOR FILING DATE: 2000-02-29
PRIOR FILING DATE: 2000-03-02
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 29
LENGTH: 535
TYPE: PRT
ORGANISM: homo sapiens
US-10-364-774-29

Query Match 93.8%; Score 217.5; DB 14; Length 535;
Best Local Similarity 95.3%; Pred. No. 8.e-22;
Matches 41; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 CGGRMEDIPCSRGVGHYRKYPYKVPAGVSLARLNKRKRVAD-WM 42
Db 288 CGGRMEDIPCSRGVGHYRKYPYKVPAGVSLARLNKRKRVAD-WM 330

RESULT 6

US-09-795-926-43
Sequence 43, Application US/09795926
Patent No. US20020098486A1

GENERAL INFORMATION:
APPLICANT: Donoho, Gregory
APPLICANT: Hilbun, Erin
APPLICANT: Turner, C. Alexander Jr.
APPLICANT: Friedrich, Glenn
APPLICANT: Abuin, Alejandro
APPLICANT: Zambrowicz, Brian
APPLICANT: Sands, Arthur T.
APPLICANT: Walke, D. Wade
APPLICANT: Kieke, James Alvin
APPLICANT: Potter, David George
TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
FILE REFERENCE: LEX-0144-USA
CURRENT FILING DATE: 2001-02-28
PRIOR FILING DATE: 2000-02-29
PRIOR FILING DATE: 2000-03-02
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 43
LENGTH: 603
TYPE: PRT
ORGANISM: homo sapiens
US-09-795-926-43

Query Match 93.8%; Score 217.5; DB 9; Length 603;
Best Local Similarity 95.3%; Pred. No. 1.e-21;
Matches 41; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 CGGRMEDIPCSRGVGHYRKYPYKVPAGVSLARLNKRKRVAD-WM 42
Db 356 CGGRMEDIPCSRGVGHYRKYPYKVPAGVSLARLNKRKRVAD-WM 398

RESULT 7

US-10-001-851-2
Sequence 42, Application US/10001851
Publication No. US20020115628A1
GENERAL INFORMATION:
APPLICANT: MEYERS, Rachel A.
APPLICANT: WILLIAMSON, Mark
TITLE OF INVENTION: 47169 and 33935, No. US20020115628A1 Human Glycosyl Transferase
FILE REFERENCE: 10147-56U1
CURRENT FILING DATE: 2001-11-20
PRIOR FILING DATE: 2000-11-20
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 603
TYPE: PRT
ORGANISM: Homo sapiens
US-10-001-851-2

Query Match 93.8%; Score 217.5; DB 13; Length 603;
Best Local Similarity 95.3%; Pred. No. 1.e-21;
Matches 41; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 CGGRMEDIPCSRGVGHYRKYPYKVPAGVSLARLNKRKRVAD-WM 42
Db 356 CGGRMEDIPCSRGVGHYRKYPYKVPAGVSLARLNKRKRVAD-WM 398

RESULT 8

US-10-364-774-43
Sequence 43, Application US/10364774
Publication No. US20030144497A1
GENERAL INFORMATION:
APPLICANT: Donoho, Gregory
APPLICANT: Hilbun, Erin
APPLICANT: Turner, C. Alexander Jr.
APPLICANT: Friedrich, Glenn
APPLICANT: Abuin, Alejandro
APPLICANT: Zambrowicz, Brian
APPLICANT: Sands, Arthur T.
APPLICANT: Walke, D. Wade
APPLICANT: Wilganowski, Nathaniel L.
APPLICANT: Hu, Yi
APPLICANT: Kieke, James Alvin
APPLICANT: Potter, David George
TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
FILE REFERENCE: LEX-0144-USA
CURRENT FILING DATE: 2003-02-11
PRIOR FILING DATE: 2001-02-28
PRIOR FILING DATE: 2000-03-02
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 43
LENGTH: 603
TYPE: PRT
ORGANISM: homo sapiens
US-10-364-774-43

Query Match 93.8%; Score 217.5; DB 14; Length 603;
Best Local Similarity 95.3%; Pred. No. 1e-21;
Matches 41; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 CGRMEDIPCSRGVGHYRKVPYKVPAGVSLARNLKRVD-WM 42
|||||
DB 356 CGRMEDIPCSRGVGHYRKVPYKVPAGVSLARNLKRVAEWM 398
|||||

RESULT 9

US-09-795-926-41

; Sequence 41, Application US/09795926
; Patent No. US20020098486A1

; GENERAL INFORMATION:

; APPLICANT: Donoho, Gregory

; APPLICANT: Hilbun, Brin

; APPLICANT: Turner, C. Alexander Jr.

; APPLICANT: Friedrich, Glenn

; APPLICANT: Abuin, Alejandro

; APPLICANT: Zambrowicz, Brian

; APPLICANT: Sands, Arthur T.

; APPLICANT: Walke, D. Wade

; APPLICANT: Wilganowski, Nathaniel L.

; APPLICANT: Hu, Yi

; APPLICANT: Kieke, James Alvin

; APPLICANT: Potter, David George

; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND

; FILE REFERENCE: LEX-0144-USA

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME

; CURRENT APPLICATION NUMBER: US/09795,926

; CURRENT FILING DATE: 2001-02-23

; PRIOR APPLICATION NUMBER: US 60/185,920

; PRIOR FILING DATE: 2000-02-29

; PRIOR APPLICATION NUMBER: US 60/186,558

; PRIOR FILING DATE: 2000-03-02

; PRIOR APPLICATION NUMBER: US 60/191,849

; PRIOR FILING DATE: 2000-03-24

; NUMBER OF SEQ ID NOS: 47

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 41

; LENGTH: 631

; TYPE: PRT

; ORGANISM: homo sapiens

US-09-795-926-41

Query Match 93.8%; Score 217.5; DB 9; Length 631;
Best Local Similarity 95.3%; Pred. No. 1e-21;
Matches 41; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 CGRMEDIPCSRGVGHYRKVPYKVPAGVSLARNLKRVD-WM 42
|||||
DB 385 CGRMEDIPCSRGVGHYRKVPYKVPAGVSLARNLKRVAEWM 427
|||||

RESULT 10

US-10-364-774-41

; Sequence 41, Application US/10364774

; Publication No. US20030144497A1

; GENERAL INFORMATION:

; APPLICANT: Donoho, Gregory

; APPLICANT: Hilbun, Brin

; APPLICANT: Turner, C. Alexander Jr.

; APPLICANT: Friedrich, Glenn

; APPLICANT: Abuin, Alejandro

; APPLICANT: Zambrowicz, Brian

; APPLICANT: Sands, Arthur T.

; APPLICANT: Walke, D. Wade

; APPLICANT: Wilganowski, Nathaniel L.

; APPLICANT: Hu, Yi

; APPLICANT: Kieke, James Alvin

; APPLICANT: Potter, David George

; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME

; FILE REFERENCE: LEX-0144-USA

; CURRENT APPLICATION NUMBER: US/10/364,774

; CURRENT FILING DATE: 2003-02-11

; PRIOR APPLICATION NUMBER: US/09/795,926

; PRIOR FILING DATE: 2001-02-28

; PRIOR APPLICATION NUMBER: US 60/185,920

; PRIOR FILING DATE: 2000-02-29

; PRIOR APPLICATION NUMBER: US 60/186,558

; PRIOR FILING DATE: 2000-03-02

; PRIOR APPLICATION NUMBER: US 60/191,849

; PRIOR FILING DATE: 2000-03-24

; NUMBER OF SEQ ID NOS: 47

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 41

; LENGTH: 631

; TYPE: PRT

; ORGANISM: homo sapiens

US-10-364-774-41

Query Match 93.8%; Score 217.5; DB 14; Length 631;

Best Local Similarity 95.3%; Pred. No. 1e-21;

Matches 41; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 CGRMEDIPCSRGVGHYRKVPYKVPAGVSLARNLKRVD-WM 42
|||||
DB 385 CGRMEDIPCSRGVGHYRKVPYKVPAGVSLARNLKRVAEWM 427
|||||

RESULT 11

US-10-292-896-3

; Sequence 3, Application US/10292896

; Publication No. US20030186850A1

; GENERAL INFORMATION:

; APPLICANT: HASSAN, Helle

; APPLICANT: REIS, Celso A.

; APPLICANT: BENNETT, Eric P.

; APPLICANT: CLAUSEN, Henrik

; TITLE OF INVENTION: METHODS TO IDENTIFY AGENTS MODULATING FUNCTIONS OF POLYPEPTIDE GAI

; TITLE OF INVENTION: TRANSFERASES, PHARMACEUTICAL COMPOSITIONS COMPRISING SUCH AGENTS

; TITLE OF INVENTION: SUCH AGENTS FOR PREPARING MEDICAMENTS

; FILE REFERENCE: 4305/1H154-US3

; CURRENT APPLICATION NUMBER: US/10/292,896

; CURRENT FILING DATE: 2003-03-17

; PRIOR APPLICATION NUMBER: US 60/425,204

; PRIOR FILING DATE: 2002-11-08

; PRIOR APPLICATION NUMBER: PCT/DK01/00328

; PRIOR FILING DATE: 2001-05-10

; PRIOR APPLICATION NUMBER: US 60/203,331

; PRIOR FILING DATE: 2000-05-11

; NUMBER OF SEQ ID NOS: 130

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 3

; LENGTH: 276

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-292-896-3

Query Match 86.4%; Score 200.5; DB 14; Length 276;

Best Local Similarity 88.9%; Pred. No. 1.1e-19;

Matches 40; Conservative 1; Mismatches 1; Indels 3; Gaps 2;

QY 1 CGRMEDIPCSRGVGHYRKVPYKVPAGVSLARNLKRVD-WM 42
|||||
DB 27 CGRMEDIPCSRGVGHYRKVPYKVPAGVSLARNLKRVAEWM 71
|||||

RESULT 12

US-09-924-358-35

; Sequence 35, Application US/09924358

; Patent No. US20020107376A1

; GENERAL INFORMATION:

; APPLICANT: Millennium Pharmaceuticals, Inc.


```

; APPLICANT: Meyers, Rachel
; APPLICANT: Macbeth, Kyle
; TITLE OF INVENTION: 26199, 33530, 33949, 47148, 50226, AND
; TITLE OF INVENTION: 58764,
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREFOR
; FILE REFERENCE: 38155-20034.00
; CURRENT APPLICATION NUMBER: US/09/924,358
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/229,300
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid
US-09-924-358-35

Query Match          59.1%; Score 137; DB 9; Length 62;
Best Local Similarity 57.8%; Pred. No. 1.8e-11;
Matches 26; Conservative 8; Mismatches 7; Indels 4; Gaps 2;

QY 1 CGGRMEDIPCSRGVGHYKVPYKVPAGVS---LARNLKRVD-W 41
Db 18 CGGKLEIVPCSRGVGHIFRKQSPYTFPGSSGNVISRNYKRVAEVW 62

RESULT 13
US-10-133-709-9
; Sequence 9, Application US/10133709
; Publication No. US20030181343A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Rosenfeld, Julie Beth
; TITLE OF INVENTION: 47133, A HUMAN GLYCOSYLTRANSFERASE
; FILE REFERENCE: 381552005200
; CURRENT APPLICATION NUMBER: US/10/133,709
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: 60/279,647
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-10-133-709-9

Query Match          59.1%; Score 137; DB 14; Length 62;
Best Local Similarity 57.8%; Pred. No. 1.8e-11;
Matches 26; Conservative 8; Mismatches 7; Indels 4; Gaps 2;

QY 1 CGGRMEDIPCSRGVGHYKVPYKVPAGVS---LARNLKRVD-W 41
Db 18 CGGKLEIVPCSRGVGHIFRKQSPYTFPGSSGNVISRNYKRVAEVW 62

RESULT 14
US-10-410-764-35
; Sequence 35, Application US/10410764
; Publication No. US2004005664A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Macbeth, Kyle J.
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Weich, Nadine S.

```

```

; APPLICANT: Olandt, Peter J.
; APPLICANT: Tsai, Fong-Ying
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Carroll, Joseph M.
; TITLE OF INVENTION: 26199, 33530, 33949, 47148, 50226,
; TITLE OF INVENTION: 58764, 62113, 32144, 32235, 23565, 13305, 14911, 86216,
; TITLE OF INVENTION: 25206 AND 8843 MOLECULES AND USES THEREFOR
; FILE REFERENCE: MPI03-0520MINM
; CURRENT APPLICATION NUMBER: US/10/410,764
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: US 09/924,358
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/229,300
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 10/350,553
; PRIOR FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: US 60/351,572
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: US 09/966,614
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/238,054
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: US 10/281,094
; PRIOR FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: US 60/347,815
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 10/076,535
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/269,440
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 136
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Amino acid consensus sequence
US-10-410-764-35

Query Match          59.1%; Score 137; DB 15; Length 62;
Best Local Similarity 57.8%; Pred. No. 1.8e-11;
Matches 26; Conservative 8; Mismatches 7; Indels 4; Gaps 2;

QY 1 CGGRMEDIPCSRGVGHYKVPYKVPAGVS---LARNLKRVD-W 41
Db 18 CGGKLEIVPCSRGVGHIFRKQSPYTFPGSSGNVISRNYKRVAEVW 62

RESULT 15
US-10-087-192-1989
; Sequence 1989, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Eggelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1989
; LENGTH: 539
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-087-192-1989

```

Query Match 55.4%; Score 128.5; DB 13; Length 539;
Best Local Similarity 55.6%; Pred. No. 3e-09;
Matches 25; Conservative 6; Mismatches 11; Indels 3; Gaps 2;
QY 1 CGGRMEDIPCSRYGHIYRKVPYKVP--AGVSLARNLKEVAD-WM 42
Db 326 CGSLEIIPCSRVGHVFRKHQHYTFEGSGTVPFARTRAAEVWM 370

Search completed: November 11, 2004, 01:28:14
Job time : 46.1692 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 12:29:32 ; Search time 9.35751 Seconds
(without alignments)
431.857 Million cell updates/sec

Title: US-10-092-750-32

Perfect score: 232

Sequence: 1 CGGRMEDIPCSRVGHIYRKVY.....YKVPAGVSLARNLKRVADWM 42

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 segs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79.*

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------|--------------------|
| 1 | 200.5 | 86.4 | 276 | T12552 | hypothetical prote |
| 2 | 120.5 | 55.4 | 571 | I37405 | polypeptide N-acet |
| 3 | 120.5 | 51.9 | 684 | T25930 | hypothetical prote |
| 4 | 116 | 50.0 | 562 | T42250 | polypeptide N-acet |
| 5 | 116 | 50.0 | 617 | T42249 | polypeptide N-acet |
| 6 | 116 | 50.0 | 618 | T42248 | polypeptide N-acet |
| 7 | 115.5 | 49.8 | 589 | T42244 | probable polypepti |
| 8 | 115 | 49.6 | 623 | T42245 | probable polypepti |
| 9 | 115 | 49.6 | 624 | T42247 | polypeptide N-acet |
| 10 | 115 | 49.6 | 626 | T42246 | polypeptide N-acet |
| 11 | 114.5 | 49.4 | 563 | A88515 | polypeptide N-acet |
| 12 | 114.5 | 49.4 | 612 | T42243 | probable polypepti |
| 13 | 109.5 | 47.2 | 559 | JC4223 | polypeptide N-acet |
| 14 | 109.5 | 47.2 | 559 | A45987 | polypeptide N-acet |
| 15 | 108 | 46.6 | 633 | JC5247 | polypeptide N-acet |
| 16 | 94 | 40.5 | 601 | T42251 | polypeptide N-acet |
| 17 | 94 | 40.5 | 605 | T27397 | hypothetical prote |
| 18 | 93 | 40.1 | 421 | T42252 | polypeptide N-acet |
| 19 | 90.5 | 39.0 | 579 | T31549 | polypeptide N-acet |
| 20 | 57.5 | 24.8 | 520 | A37844 | hypothetical prote |
| 21 | 57 | 24.6 | 351 | C64646 | dihydroorotate deh |
| 22 | 57 | 24.6 | 540 | T34702 | hypothetical prote |
| 23 | 57 | 24.6 | 1027 | T38604 | p53-binding protei |
| 24 | 56.5 | 24.4 | 792 | T50122 | vesicular transpor |
| 25 | 56 | 24.1 | 536 | A39036 | H+-transporting tw |
| 26 | 55 | 23.7 | 305 | D64326 | UDPglucose 4-epime |
| 27 | 55 | 23.7 | 351 | E71935 | dihydroorotate deh |
| 28 | 55 | 23.7 | 374 | S41758 | heat shock protein |
| 29 | 55 | 23.7 | 374 | C97058 | molecular chaperon |

RESULT 1
T12552
hypothetical protein DKFZps86H0623.1 - human (fragments)
C:Species: Homo sapiens (man)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 09-Jul-2004
C:Accession: T12552
R:Ansoerge, W.; Witkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, June 1999
A:Reference number: 217527
A:Accession: T12552
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-150;151-276 <ANS>
A:Cross-references: UNIPROT:Q9Y4M4; EMBL:AL096739
A:Experimental source: adult uterus; clone DKFZps86H0623
A>Note: the cDNA sequence contains a -1 frameshift near codon 150
C:Genetics:
A>Note: DKFZps86H0623.1
C:Superfamily: polypeptide N-acetylglucosaminyltransferase

Query Match 86.4%; Score 200.5; DB 2; Length 276;
Best Local Similarity 88.9%; Pred. No. 3.8e-19;
Matches 40; Conservative 1; Mismatches 1; Indels 3; Gaps 2;

QY 1 CGGRMEDIPCSRVGHIYRKVVPYKVPAGVSLA--RNLKRVAD-WM 42
|||||
Db 27 CGGRMEDIPCSRVGHIYRKVVPYKVPAGVSLARVRLKRVAEVWM 71
|||||

RESULT 2
I37405
polypeptide N-acetylglucosaminyltransferase (EC 2.4.1.41) - human
C:Species: Homo sapiens (man)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 09-Jul-2004
C:Accession: I37405
R:White, T.; Bennett, E.P.; Paul, E.; Takio, K.; Srensen, T.; Bonding, N.; Clausen, H.
J. Biol. Chem. 270 24156-24165, 1995
A:Title: Purification and cDNA cloning of a human UDP-GalNAC:polypeptide N-Acetylglucosaminyltransferase
A:Reference number: I37404; MUID:96025800; PMID:7592619
A:Accession: I37405
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-571 <RES>
A:Cross-references: UNIPROT:Q10471; EMBL:X85019; NID:g971460; PIDN:CAA59381.1; PID:g97146
C:Genetics:
A:Gene: GDB:GALNT2; GALNAC
A:Cross-references: GDB:696223
A:Map position: 16q24-16q24
C:Superfamily: polypeptide N-acetylglucosaminyltransferase
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 55.4%; Score 128.5; DB 2; Length 571;

30 55 23.7 447 2 S37048
31 54.5 23.5 360 2 T20853
32 54.5 23.5 498 2 T06667
33 54.5 23.5 499 2 G96831
34 53.5 23.1 386 2 T16010
35 53.5 23.1 390 2 B72639
36 53.5 23.1 1408 2 T45039
37 53 22.8 304 2 T25337
38 53 22.8 553 2 C83444
39 53 22.8 602 2 A89868
40 52.5 22.6 180 2 F75607
41 52.5 22.6 209 2 C87530
42 52 22.4 313 2 S61990
43 52 22.4 761 2 E64449
44 51.5 22.2 509 2 T51806
45 51.5 22.2 774 1 JQ0550

ALIGNMENTS

cysteine proteinas
hypothetical prote
argininosuccinate
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
probable AMP-bind
hypothetical prote
arsenate reductase
uracil phosphoribo
hypothetical prote
hypothetical prote
cytochrome P450 ka
1,4-alpha-glucan b

R;Hagen, F.K.; Nehrke, K.
J. Biol. Chem. 273, 8268-8277, 1998
A:Title: cDNA cloning and expression of a family of UDP-N-acetyl-Dgalactosamine:Polypeptidyl transferase C superfamily members from *Caenorhabditis elegans*
A:Reference number: Z22126; MUID:98192620; PMID:9525933
A:Accession: T42249
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-617 <HAG>
A:Cross-references: UNIPROT:O61395; EMBL:AF031839; NID:g3047198; PIDN:AAC13675.1; PID:g3047196; CESP:H38K22.5
C:Genetics:
A:Gene: gly-6
C:Superfamily: polypeptide N-acetyl-galactosaminyltransferase
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 50.0%; Score 116; DB 2; Length 617;
Best Local Similarity 52.2%; Pred. No. 1.7e-07;
Matches 24; Conservative 7; Mismatches 11; Indels 4; Gaps 2;

QY 1 CGGRMEDPCSRVGHYKYVPYKP---AGVSLARNLKRVAD-WM 42
|||||:|||||||:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 372 CGGRVEILPCSHGVHFRKSPHPDPGKSGKVLNTLLRVAEVMW 417

RESULT 5
T42248
polypeptide N-acetylgalactosaminyltransferase (EC 2.4.1.41) - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T42248; T23138
R;Hagen, F.K.; Nehrke, K.
J. Biol. Chem. 273, 8268-8277, 1998
A:Title: cDNA cloning and expression of a family of UDP-N-acetyl-Dgalactosamine:Polypeptidyl transferase C superfamily members from *Caenorhabditis elegans*
A:Reference number: Z22126; MUID:98192620; PMID:9525933
A:Accession: T42248
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-618 <HAG>
A:Cross-references: UNIPROT:O61394; EMBL:AF031838; NID:g3047196; PIDN:AAC13674.1; PID:g3047196; CESP:H38K22.5
C:Genetics:
A:Gene: H38K22.5
A:Map position: 3
A:Introns: 42/3; 64/3; 154/1; 202/1; 368/3; 430/1; 477/3; 513/3; 563/3; 604/3
A>Note: gly-6
C:Superfamily: polypeptide N-acetyl-galactosaminyltransferase
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 50.0%; Score 116; DB 2; Length 618;
Best Local Similarity 52.2%; Pred. No. 1.7e-07;
Matches 24; Conservative 7; Mismatches 11; Indels 4; Gaps 2;

QY 1 CGGRMEDPCSRVGHYKYVPYKP---AGVSLARNLKRVAD-WM 42
|||||:|||||||:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 372 CGGRVEILPCSHGVHFRKSPHPDPGKSGKVLNTLLRVAEVMW 417

RESULT 7
T42244
probable polypeptide N-acetyl-galactosaminyltransferase (EC 2.4.1.41) - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T42244
R;Hagen, F.K.; Nehrke, K.
J. Biol. Chem. 273, 8268-8277, 1998
A:Title: cDNA cloning and expression of a family of UDP-N-acetyl-Dgalactosamine:Polypeptidyl transferase C superfamily members from *Caenorhabditis elegans*

A;Title: cDNA cloning and expression of a family of UDP-N-acet-

Search completed: November 10, 2004, 13:40:40
Job time : 10.3575 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 12:27:34 ; Search time 51.285 Seconds
(without alignments)
471.205 Million cell updates/sec

Title: US-10-092-750-32

Perfect score: 232

Sequence: 1 CGGRMEDPCSRVGHYRKY.....YKVPAGVSLARNLKKVADWM 42

Scoring table: BLOSUM62

Gapop 10.0., Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Watch 100%

Listing first 45 summaries

Database : UniProt 02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|-------------|
| 1 | 217.5 | 93.8 | 603 | 1 | GL10_HUMAN |
| 2 | 217.5 | 93.8 | 603 | 1 | GL10_MOUSE |
| 3 | 217.5 | 93.8 | 603 | 1 | GL10_RAT |
| 4 | 217.5 | 93.8 | 603 | 2 | AAH60617 |
| 5 | 217.5 | 93.8 | 634 | 2 | BAD21405 |
| 6 | 143 | 61.6 | 599 | 2 | Q7P2M5 |
| 7 | 131 | 56.5 | 644 | 2 | Q7Q0E9 |
| 8 | 128.5 | 55.4 | 558 | 1 | GLT1_HUMAN |
| 9 | 128.5 | 55.4 | 558 | 1 | GLT1_MOUSE |
| 10 | 128.5 | 55.4 | 570 | 1 | GLT2_MOUSE |
| 11 | 128.5 | 55.4 | 570 | 2 | AAH59818 |
| 12 | 128.5 | 55.4 | 571 | 1 | GLT2_HUMAN |
| 13 | 128 | 55.2 | 1003 | 2 | Q7QD80 |
| 14 | 126.5 | 54.5 | 590 | 2 | AAH64620 |
| 15 | 126.5 | 54.5 | 615 | 2 | AAQ56700 |
| 16 | 126.5 | 54.5 | 633 | 1 | GLT2_DROME |
| 17 | 126 | 54.3 | 351 | 2 | Q7Q0X9 |
| 18 | 122.5 | 52.8 | 620 | 2 | Q7Q456 |
| 19 | 120.5 | 51.9 | 622 | 1 | GL10_CAEEL |
| 20 | 118 | 50.9 | 518 | 2 | Q7Q048 |
| 21 | 118 | 50.9 | 650 | 1 | GLT3_DROME |
| 22 | 117.5 | 50.6 | 667 | 1 | GLT3_DROME |
| 23 | 116.5 | 50.2 | 550 | 1 | GL14_MOUSE |
| 24 | 116.5 | 50.2 | 552 | 1 | GL14_HUMAN |
| 25 | 116.5 | 50.2 | 552 | 2 | AAQ89118 |
| 26 | 116.5 | 50.2 | 645 | 2 | Q7Q815 |
| 27 | 116 | 50.0 | 618 | 1 | GLT8_CAEEL |
| 28 | 115.5 | 49.8 | 589 | 1 | GLT8_CAEEL |
| 29 | 115 | 49.6 | 626 | 1 | GLT5_CAEEL |
| 30 | 114.5 | 49.4 | 612 | 1 | GLT3_CAEEL |
| 31 | 114 | 49.1 | 622 | 1 | GLT6_HUMAN |

32 114 49.1 622 1 GLT6_MOUSE
33 113.5 48.9 930 1 GLT5_RAT
34 113.5 48.9 940 1 GLT5_HUMAN
35 112.5 48.5 598 1 GLT3_HUMAN
36 112.5 48.5 598 1 GLT3_MOUSE
37 112.5 48.5 598 2 AAH67524
38 112.5 48.5 598 2 AAH67525
39 112.5 48.5 598 2 AAH69624
40 112.5 48.5 598 2 AAH69628
41 112.5 48.5 598 2 AAH69636
42 112.5 48.5 598 2 AAH69645
43 112.5 48.5 598 2 AAH69997
44 112.5 48.5 930 1 GLT5_MOUSE
45 111.5 48.1 601 1 GLT1_DROME

ALIGNMENTS

RESULT 1

GL10_HUMAN STANDARD; Q86VP8; Q8IXJ2; Q96IV2; Q9H8E1; Q9Y4M4;
ID AC Q86SRL; Q6INS6; Q86VP8; Q8IXJ2; Q96IV2; Q9H8E1; Q9Y4M4;
DT 01-OCT-2004 (Rel. 45, Created)
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DE Polypeptide N-acetylgalactosaminyltransferase 10 (EC 2.4.1.41)
DE (Protein-UDP acetylglactosaminyltransferase 10) (UDP-
DE GalNAc:polypeptide N-acetylglactosaminyltransferase 10) (UDP-
DE GalNAc transferase 10) (GalNAc-T10) (pp-GalTase 10).
GN Name=GALNT10;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), ENZYME ACTIVITY, AND TISSUE
RP SPECIFICITY.
RC TISSUE=Colon cancer;
RX MEDLINE=22304871; PubMed=12417297;
RA Cheng L., Tachibana K., Zhang Y., Guo J.-M., Tachibana K.K.,
RA Kameyama A., Wang H., Hiruma T., Iwasaki H., Togayachi A., Kudo T.,
RA Narimatsu H.;
RT "Characterization of a novel human UDP-GalNAc transferase, pp-GalNAc-
RT T10.";
RL FEBS Lett. 531:115-121(2002).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Bennett E.P.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 4).
RC TISSUE=Uterus;
RA Ansong W., Wirkner U., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 5), AND SEQUENCE OF 191-603 FROM N.A.
RP (ISOFORM 1).
RC TISSUE=Kidney, Skin, and Testis;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan K., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Rosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilov D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5].
 RP SEQUENCE OF 8-366 FROM N.A. (ISOFORM 3), AND SEQUENCE OF 117-603 FROM
 RP N.A. (ISOFORM 2).
 RC TISSUE=Placenta, and Spleen;
 RX PubMed=14702039; DOI=10.1038/ng1285;
 RA Ota T., Suzuki K., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakatsuki A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Ohayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
 RA Negahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
 RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
 RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
 RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
 RA Fujimori K., Tanai H., Kimata S., Watanabe M., Hirakawa S., Chiba Y.,
 RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
 RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
 RA Nomura Y., Togiyasu S., Konai F., Hara R., Takeuchi K., Arita M.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
 RA Moriya S., Momiya Y., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takenoto M., Wakabe H.,
 RA Yamagaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Saeki M.,
 RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs";
 RL Nat. Genet. 36:40-45(2004).
 CC -!- FUNCTION: Catalyzes the initial reaction in O-linked
 CC oligosaccharide biosynthesis, the transfer of an N-acetyl-D-
 CC galactosamine residue to a serine or threonine residue on the
 CC protein receptor. Has activity toward Muc5AC and E2 peptide
 CC substrates.
 CC -!- CATALYTIC ACTIVITY: UDP-N-acetyl-D-galactosamine + polypeptide =
 CC UDP + N-acetyl-D-galactosaminyl-polypeptide.
 CC -!- COFACTOR: Manganese and calcium (By similarity).
 CC -!- PATHWAY: Glycosylation.
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (By
 CC similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=5;
 CC Name=1;
 CC IsoId=Q86SR1-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q86SR1-2; Sequence=VSP_011209;
 CC Note=No experimental confirmation available;
 CC Name=3;
 CC IsoId=Q86SR1-3; Sequence=VSP_011212, VSP_011213;
 CC Note=No experimental confirmation available;
 CC Name=4;
 CC IsoId=Q86SR1-4; Sequence=VSP_011207, VSP_011208, VSP_011214;
 CC Note=No experimental confirmation available;
 CC Name=5;
 CC IsoId=Q86SR1-5; Sequence=VSP_011210, VSP_011211;
 CC Note=No experimental confirmation available;
 CC -!- TISSUE SPECIFICITY: Widely expressed. Expressed at high level in
 CC small intestine, and at intermediate levels in stomach, pancreas,
 CC ovary, thyroid gland and spleen. Weakly expressed in other

CC tissues.
 CC -!- DOMAIN: There are two conserved domains in the glycosyltransferase
 CC region: the N-terminal domain (domain A, also called G1 motif),
 CC which is probably involved in manganese coordination and substrate
 CC binding and the C-terminal domain (domain B, also called
 CC Gal/galnac-T motif), which is probably involved in catalytic
 CC reaction and UDP-Gal binding (By similarity).
 CC -!- DOMAIN: The ricin B-type lectin domain binds to GalNAc and
 CC contributes to the glycopeptide specificity (By similarity).
 CC -!- SIMILARITY: Belongs to the glycosyltransferase family 2. Galnac-T
 CC subfamily.
 CC -!- SIMILARITY: Contains 1 ricin B-type lectin domain.
 CC -!- CAUTION: According to experiments made in rat, this enzyme is
 CC unable to transfer GalNAc onto serine or threonine residue on the
 CC protein receptor, but instead requires the prior addition of a
 CC GalNAc on a peptide before adding additional GalNAc moieties,
 CC thereby acting as a glycopeptide transferase.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB078145; BAC56890.1; -;
 CC EMBL; AJ505950; CAD44532.1; -;
 CC EMBL; AL096739; CAB46378.1; -;
 CC EMBL; BC007224; AAH07224.2; -;
 CC EMBL; BC050333; AAH50333.1; -;
 CC EMBL; BC072450; AAH72450.1; -;
 CC EMBL; AK023782; BAB14676.1; ALT_INIT.
 CC EMBL; AK074132; BAB84958.1; -;
 CC PIR; T12552; T12552.
 CC HSSP; P26514; 1XNM.
 CC Genew; HGNC:19873; GALNT10.
 CC MIM; 608043; -;
 CC InterPro; IPR001173; Glyco_trans_2.
 CC InterPro; IPR008997; RicinB-like.
 CC InterPro; IPR000772; Ricin_B_lectin.
 CC Pfam; PF00535; Glycos_transf_2; 1.
 CC Pfam; PF00652; Ricin_B_lectin; 3.
 CC SMART; SMC00458; RICIN_1.
 CC PROSITE; PS0231; RICIN_B_LECTIN; 1.
 CC Alternative splicing; Calcium; Glycosyltransferase; Golgi stack;
 CC Lectin; Manganese; Signal-anchor; Transmembrane.
 CC DOMAIN 1 11 Cytoplasmic (Potential).
 CC TRANSMEM 12 31 Signal-anchor for type II membrane
 CC protein (Potential).
 CC Lumenal (Potential).
 CC Catalytic subdomain A.
 CC Domain 311 373 Catalytic subdomain B.
 CC DOMAIN 458 590 Ricin B-type lectin.
 CC DISULFID 471 488 By similarity.
 CC DISULFID 523 538 By similarity.
 CC DISULFID 563 578 By similarity.
 CC CARBOHYD 124 124 N-linked (GlcNAc...) (Potential).
 CC CARBOHYD 146 146 N-linked (GlcNAc...) (Potential).
 CC CARBOHYD 593 593 N-linked (GlcNAc...) (Potential).
 CC VARSPPLIC 1 329 Missing (in isoform 4).
 CC FTID=VSP_011207.
 CC FT VARSPPLIC 330 352 WELGGYDPEIWEQGEYSFK -> MLAWRDGELEAETS
 CC SSLLFLAMQ (in isoform 4).
 CC FTID=VSP_011208.
 CC Missing (in isoform 2).
 CC FTID=VSP_011209.
 CC Missing (in isoform 5).
 CC EHLKKPLEDFMAL -> DLPTASTPSPFVC (in
 CC isoform 5).
 CC FTID=VSP_011210.
 CC Missing (in isoform 5).
 CC FTID=VSP_011211.
 CC WMCGRMEDTPCS -> SQLSRPVLGTAS (in
 CC isoform 5).

FT isoform 3).
 FT /FTID=VSP_011212.
 Query Match 93.8%; Score 217.5; DB 1; Length 603;
 Best Local Similarity 95.3%; Fred. No. 3.6e-21;
 Matches 41; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
 QY 1 CGGRMEDIPCSRVGHYRKYVPKYPAGVSLARNLKRVAWM 398
 DB 356 CGGRMEDIPCSRVGHYRKYVPKYPAGVSLARNLKRVAEYWM 398
 RESULT 2
 GL10 MOUSE
 ID GL10 MOUSE STANDARD; PRT; 603 AA.
 AC G6P357; Q6KQAQ2; Q8BZU8; Q91XJ6;
 DT 01-OCT-2004 (Rel. 45, Created)
 DT 01-OCT-2004 (Rel. 45, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Polypeptide N-acetylgalactosaminyltransferase 10 (EC 2.4.1.41)
 DE (Protein-UDP acetylgalactosaminyltransferase 10) (UDP-
 DE GalNac:polypeptide N-acetylgalactosaminyltransferase 10) (Polypeptide
 DE GalNac transferase 10) (GalNac-T10) (pp-GaNTase 10).
 GN Name-GaNT10;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
 RA Suga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.,
 RA Klausner R.D., Collins F.S., Wagner K.H., Shermen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Murzyn D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences".;
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [3]
 RP SEQUENCE OF 76-603 FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Colon;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klapper S.L., Collins L., Culbertson L.M., Chen L.M., Niggli V.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Murzyn D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences".;
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [3]

Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
 Konoaga A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 Nagashima T., Numata K., Okido T., Pavan W.J., Perlea G., Pesole G.,
 Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 Sardinia A., Schneider C., Semple C.A., Setou M., Shimada K.,
 Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 Varadar L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang C.,
 Wilming L.G., Wilming L., Wilming L., Wilming L., Wilming L.,
 Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 Birney E., Hayashizaki Y.,
 "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs".;
 Nature 420:563-573 (2002).
 [4]
 TISSUE SPECIFICITY.
 MEDLINE=21264503; PubMed=11278534; DOI=10.1074/jbc.M009638200;
 RA Ten Hagen K.G., Bedi G.S., Tetaert D., Kingsley P.D., Hagen F.K.,
 RA Balys M.M., Beres T.M., Degand P., Tabak L.A.,
 RA "Cloning and characterization of a ninth member of the UDP-
 RT GalNac:polypeptide N-acetylgalactosaminyltransferase family,
 RT ppGaNTase-T9".;
 RL J. Biol. Chem. 276:17395-17404 (2001).
 CC -!- FUNCTION: Catalyzes the initial reaction in O-linked
 CC oligosaccharide biosynthesis, the transfer of an N-acetyl-D-
 CC galactosamine residue to a serine or threonine residue on the
 CC protein receptor. Has activity toward Muc5AC and EA2 peptide
 CC substrates (By similarity).
 CC -!- CATALYTIC ACTIVITY: UDP-N-acetyl-D-galactosamine + polypeptide =
 CC UDP + N-acetyl-D-galactosaminyl-polypeptide.
 CC -!- COFACTOR: Manganese and calcium (By similarity).
 CC -!- PATHWAY: Glycosylation.
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (By
 CC similarity).
 CC -!- TISSUE SPECIFICITY: Expressed at higher level than GALNT9. In the
 CC developing hindbrain region of E14.5 embryos it accumulates in the
 CC rapidly dividing, undifferentiated ventricular zone adjacent to the
 CC pons. It also accumulates in the regions immediately rostral
 CC and caudal to the dorsal rhombic lips differentiating into the
 CC cerebellum. Not expressed in the developing choroid plexus.
 CC -!- DOMAIN: There are two conserved domains in the glycosyltransferase
 CC region: the N-terminal domain (domain A, also called G1 motif),
 CC which is probably involved in manganese coordination and substrate
 CC binding and the C-terminal domain (domain B, also called
 CC Gal/GalNac-T motif), which is probably involved in catalytic
 CC reaction and UDP-Gal binding (By similarity).
 CC -!- DOMAIN: The ricin B-type lectin domain binds to GalNac and
 CC contributes to the glycopeptide specificity (By similarity).
 CC -!- SIMILARITY: Belongs to the glycosyltransferase family 2. GalNac-T
 CC subfamily.
 CC -!- SIMILARITY: Contains 1 ricin B-type lectin domain.
 CC -!- CAUTION: According to experiments made in rat, this enzyme is
 CC unable to transfer GalNac onto serine or threonine residue on the
 CC protein receptor, but instead requires the prior addition of a
 CC GalNac on a peptide before adding additional GalNac moieties,
 CC thereby acting as a glycopeptide transferase.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; AK131155; BAD21405.1; ALT_INIT.

DR EMBL; BC016585; AAH16585.1; ALT_INIT.
 DR EMBL; BC060617; AAH60617.1; -
 DR EMBL; AK033515; BAC28334.1; -
 DR HSP; P28514; IKNL.
 DR MGD; MGI:1890480; Galnt10.
 DR GO; GO:0004653; F-polypeptide N-acetylgalactosaminyltransferase. . . ; IDA.
 DR GO; GO:0006493; P-O-linked glycosylation; IDA.
 DR InterPro; IPR001173; Glyco trans 2.
 DR InterPro; IPR008997; RicinB-like.
 DR InterPro; IPR000772; RicinB-lectin.
 DR Pfam; PF00535; Glycos transf 2; 1.
 DR Pfam; PF00652; RicinB-lectin; 2.
 DR SMART; SM00458; RICIN; 1.
 DR PROSITE; PS02031; RICIN_B-LECTIN; 1.
 KW Calcium; Glycosyltransferase; Golgi stack; Lectin; Manganese;
 KW Signal-anchor; Transferrase; Transmembrane.
 FT DOMAIN 1 11
 FT TRANSMEM 12 31
 FT
 FT DOMAIN 32 603
 FT DOMAIN 144 253
 FT DOMAIN 311 373
 FT DOMAIN 458 590
 FT DISULFID 471 488
 FT DISULFID 523 538
 FT DISULFID 563 578
 FT CARBOHYD 124 124
 FT CARBOHYD 146 146
 FT CARBOHYD 593 593
 FT CONFLICT 233 233
 FT CONFLICT V -> I (in Ref. 1).
 SQ SEQUENCE 603 AA; 69116 MW; PF55FBA7E1DD7544 CRC64;

Query Match 93.8%; Score 217.5; DB 1; Length 603;
 Best Local Similarity 95.3%; Pred. No. 3.6e-21;
 Matches 41; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 CGRMEDIPCSRVGHYKRVKVPKPVAGVSLARNLRKRVAD-WM 42
 DB 356 CGRMEDIPCSRVGHYKRVKVPKPVAGVSLARNLRKRVAD-WM 398

RESULT 3
 ID_GL10_RAT STANDARD; PRT; 603 AA.
 AC Q22557;
 DT 01-OCT-2004 (Rel. 45, Created)
 DT 01-OCT-2004 (Rel. 45, Last sequence update)
 DE Polypeptide N-acetylgalactosaminyltransferase 10 (EC 2.4.1.41)
 DE (Protein-UDP acetylgalactosaminyltransferase 10) (UDP-
 DE GalNAc-polypeptide N-acetylgalactosaminyltransferase 10) (Polypeptide
 DE GalNAc transferase 10) (GalNAc-T10) (pp-ganTase 10).
 GN Name=Galnt10;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]_TaxID=10116;
 RP SEQUENCE FROM N.A., ENZYME ACTIVITY, AND TISSUE SPECIFICITY.
 RC TISSUE=Sublingual gland;
 RX MEDLINE=21264503; PubMed=11278534; DOI=10.1074/jbc.M009638200;
 RA Ten Hagen K.G., Bedi G.S., Tetaert D., Kingaleij P.D., Hagen F.K.,
 RA Balyas M.M., Beres T.M., Degan P., Tabak L.A.;
 RT "Cloning and characterization of a ninth member of the UDP-
 RT GalNAc-polypeptide N-acetylgalactosaminyltransferase family,
 RT ppGanTase-10".
 RL J. Biol. Chem. 276:17395-17404 (2001).
 CC -!- FUNCTION: Catalyzes the initial reaction in O-linked
 CC oligosaccharide biosynthesis, the transfer of an N-acetyl-D-
 CC galactosamine residue to a serine or threonine residue on the
 CC protein receptor. Has activity toward Muc5Ac and EA2 peptide
 CC substrates.
 CC -!- CATALYTIC ACTIVITY: UDP-N-acetyl-D-galactosamine + polypeptide =

CC UDP + N-acetyl-D-galactosaminyl-polypeptide.
 CC -!- PATHWAY: Manganese and calcium (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (By
 CC similarity).
 CC -!- TISSUE SPECIFICITY: Highly expressed in the sublingual gland,
 CC testis, small intestine, colon and ovary. Expressed at
 CC intermediate level in heart, brain, spleen, lung, stomach, cervix
 CC and uterus.
 CC -!- DOMAIN: There are two conserved domains in the glycosyltransferase
 CC region: the N-terminal domain (domain A, also called G1 motif),
 CC which is probably involved in manganese coordination and substrate
 CC binding and the C-terminal domain (domain B, also called
 CC Gal/GalNAc-T motif), which is probably involved in catalytic
 CC reaction and UDP-Gal binding (By similarity).
 CC -!- DOMAIN: The ricin B-type lectin domain binds to GalNAc and
 CC contributes to the glycopeptide specificity (By similarity).
 CC -!- SIMILARITY: Belongs to the glycosyltransferase family 2. GalNAc-T
 CC subfamily.
 CC -!- SIMILARITY: Contains 1 ricin B-type lectin domain.
 CC -!- CAUTION: According Ref.1, this enzyme is unable to transfer GalNAc
 CC onto serine or threonine residue on the protein receptor, but
 CC instead requires the prior addition of a GalNAc on a peptide
 CC before adding additional GalNAc moieties, thereby acting as a
 CC glycopeptide transferase.
 CC -!- CAUTION: Was originally (Ref.1) termed Galnt9/pp-GanTase 9.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF241241; AAK54498.1; -
 CC HSP; P28514; IKNL.
 CC RGD; 6409; Galnt10.
 CC InterPro; IPR001173; Glyco trans 2.
 CC InterPro; IPR008997; RicinB-like.
 CC InterPro; IPR000772; RicinB-lectin.
 CC Pfam; PF00535; Glycos transf 2; 1.
 CC Pfam; PF00652; RicinB-lectin; 3.
 CC SMART; SM00458; RICIN; 1.
 CC PROSITE; PS02031; RICIN_B-LECTIN; 1.
 KW Calcium; Glycosyltransferase; Golgi stack; Lectin; Manganese;
 KW Signal-anchor; Transferrase; Transmembrane.
 FT DOMAIN 1 11
 FT TRANSMEM 12 31
 FT
 FT DOMAIN 32 603
 FT DOMAIN 144 253
 FT DOMAIN 311 373
 FT DOMAIN 458 590
 FT DISULFID 471 488
 FT DISULFID 523 538
 FT DISULFID 563 578
 FT CARBOHYD 124 124
 FT CARBOHYD 146 146
 FT CARBOHYD 593 593
 SQ SEQUENCE 603 AA; 194EDEAA626A4BBF CRC64;
 Query Match 93.8%; Score 217.5; DB 1; Length 603;
 Best Local Similarity 95.3%; Pred. No. 3.6e-21;
 Matches 41; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
 QY 1 CGRMEDIPCSRVGHYKRVKVPKPVAGVSLARNLRKRVAD-WM 42
 DB 356 CGRMEDIPCSRVGHYKRVKVPKPVAGVSLARNLRKRVAD-WM 398
 RESULT 4
 AAH60617

```

ID AAH60617 PRELIMINARY; PRT; 603 AA.
AC AAH60617;
DT 12-MAY-2004 (TrEMBLrel. 27, Created)
DT 12-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DT 12-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE UDP-N-acetyl-alpha-D-galactosamine:polypeptide
DE N-acetyl-galactosaminyltransferase 10.
DE GALNT10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feilings F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Klausner R.D., Collins F.S., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Wang J., Hsieh F.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX Strausberg R.;
RA Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC060617; AAH60617.1; -.
RW Transferase.
SQ SEQUENCE 603 AA; 69116 MW; FF55FBA7E1DD7544 CRC64;

Query Match 93.8%; Score 217.5; DB 2; Length 603;
Best Local Similarity 95.3%; Pred. No. 3.6e-21;
Matches 41; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 CGGRMEDIPCSRUGHYIKYVPYKVPAGVSLARNLKRVAAD-WM 42
DB 356 CGGRMEDIPCSRUGHYIKYVPYKVPAGVSLARNLKRVAEVMW 398

RESULT 5
BAD21405 PRELIMINARY; PRT; 634 AA.
ID BAD21405;
AC BAD21405;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE MFLJ00205 protein (Fragment).
DE MFLJ00205.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Suga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of FLJ Gene:
RT The Complete Nucleotide Sequences of 110 Mouse FLJ-homologous cDNAs

```

```

RT Identified by Screening of Terminal sequences of cDNA Clones Randomly
RT Sampled from Size-Fractionated Libraries";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AKJ31155; BAD21405.1; -.
FT NON TER
SQ SEQUENCE 634 AA; 72315 MW; C79862D4D0052C81 CRC64;

Query Match 93.8%; Score 217.5; DB 2; Length 634;
Best Local Similarity 95.3%; Pred. No. 3.8e-21;
Matches 41; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 CGGRMEDIPCSRUGHYIKYVPYKVPAGVSLARNLKRVAAD-WM 42
DB 387 CGGRMEDIPCSRUGHYIKYVPYKVPAGVSLARNLKRVAEVMW 429

RESULT 6
Q7PZM5 PRELIMINARY; PRT; 599 AA.
ID Q7PZM5;
AC Q7PZM5;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AGCP9340 (Fragment).
GN Name=agCG54449; ORFNames=ENSANGG00000012813;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB01008986; EAA00339.1; -.
DR InterPro; IPR001173; Glyco_trans_2.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF00535; Glycos_transf_1.
DR Pfam; PF00652; Ricin_B_lectin; 3.
DR PROSITE; PS50231; RICIN_B_LECTIN; 1.
FT NON TER
SQ SEQUENCE 599 AA; 69224 MW; DA646C182B143028 CRC64;

Query Match 61.6%; Score 143; DB 2; Length 599;
Best Local Similarity 61.4%; Pred. No. 7.4e-11;
Matches 27; Conservative 6; Mismatches 9; Indels 2; Gaps 2;

QY 1 CGGRMEDIPCSRUGHYIKYVPYKVPAGVSLARNLKRVAAD-WM 42
DB 348 CGGRMEDIPCSRUGHYIKYVPYKVPAGVSLARNLKRVAEVMW 391

RESULT 7
Q7Q0E9 PRELIMINARY; PRT; 644 AA.
ID Q7Q0E9;
AC Q7Q0E9;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AGCP9480 (Fragment).
GN Name=agCG54007; ORFNames=ENSANGG00000009226;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

```

